Canonical Perturbative Approach to Strongly Correlated Systems

By
Raymond YungKai Chan

30 September, 2004
First Corrected version : 15 March 2005
Second corrected version : 15 July 2005

A thesis submitted for the degree of Doctor of Philosophy
of the Australian National University
Declaration

I certify that the work contained in this thesis is my own original research, produced in collaboration with my supervisor – Doctor M. Gulácsi. All material taken from other references is explicitly acknowledged as such. I also certify that the work contained in this thesis has not been submitted for any other degree.

Raymond Chan
Acknowledgments

It is a pleasure to thank the people who have encouraged and supported me during the course of my PhD research, and finally the production of this thesis.

The first one I would like to thank is my supervisor Doctor Miklós Gulácsi who showed me the new world of condensed matter physics seven years ago and inspired me over the years the fascinations of this field of physics. His patience and tolerance were very much appreciated, especially during the difficult times when the progress was extremely slow and the research was undertaken on a part-time basis due to my financial reason.

The second person I would like to thank is my wife Christine who helped me in many different ways to the finish of this thesis. Her warmth and love, together with her patience and understanding, made the difficult task enjoyable. I am also very grateful to her patience of proof-reading this thesis, as her knowledge in this field of physics was very little.

I would like to thank my father, my mother and my sister for their love and support, and their help in doing some of the works around the house for me, so that I could concentrate on finishing the research. My dogs, Darcy and Coacoa, also spent a lot of time playing with me when I needed a break from the research, which I appreciated greatly.

Finally, I would like to thank the employee of CEA for their support and encouragement.
Abstract

Despite the impressive advance of materials science and technologies in high temperature superconductors, the theoretical progress of these superconductors are still very slow. Various models have been proposed in the past towards the interactions of electrons or holes in these superconductive materials, but none of them is able to provide a satisfactory explanation for the collective phenomenon of superconductivity. Amongst all these different type of models, the periodic Anderson model is the most favourable one which is believed to provide the best description for the behaviour of electrons or holes in the atomic level. Our understanding of the model is however very limited. The periodic Anderson model is a lattice extension of the single impurity Anderson model proposed initially by P.W. Anderson [6] in 1961 to explain the ferromagnetic nature of alloys in the mean field approximation. Although the results of the mean field calculation were short lived, the model itself has been widely studied and applied to many different materials since then. A large amount of literature were devoted to the study of this single impurity model and abundant theories were established for it. Even exact solution is available for the 1/2 spin model. These results of the single impurity model nonetheless cannot be generalized to the lattice case and even now the periodic Anderson model still remains to be mysterious to a certain extent.

The objective of this thesis is to study the periodic Anderson model (PAM) using canonical perturbation and canonical transformation. Some believed that perturbative approach is bound to fail because of the weak interaction approximation in a strongly interacting system. It has however been shown in many cases that perturbation theory can yield the correct behaviour of systems even in the commonly accepted non-perturbative regime, provided the perturbed parameters are chosen properly.

In this thesis, we have provided the perturbative calculation of the lattice model in three different limits, the small $U$ limit, the atomic limit and the weak hopping limit. New results were found in spin interaction among the electrons in the atomic
limit and the weak hopping limit. The same model was also investigated using a canonical transformation, similar to the one Schrieffer and Wolff [78] used in 1966 to establish a relationship between the Anderson and the Kondo Hamiltonians. Contrary to the latter transformation, we carried out the transformation exactly to infinity order. Higher order terms generated more new terms, but the low-energy contributions of these new terms were not all distinct. The low-energy results of the transformation to any order were summarized in a form in which each term of the result has a unique set of coefficients. The complexity of the coefficients increases with increasing order of transformation. There was however a pattern among the coefficients of different orders of the transformation which allowed us to generalize them to any order. Using induction, we were able to write down the transformation to all orders and subsequently summed them together to a simple form. This transformation, as set out by Schrieffer and Wolff initially, removed the first order hybridization process but higher order hybridization still remained. Other forms of transformations were examined towards the end of the thesis, to remove the entire hybridization process from the Hamiltonian.

In summary, this thesis includes examination of the PAM in the small $U$ limit, the weak hopping limit and the atomic limit using perturbation theory in chapters 2, 3 and 4. Diagram techniques were used in the atomic limit to find the sum of two perturbative expansions which yielded new interesting results. Chapters 6 and 7 contain the major work of this thesis which include the canonical transformation of the PAM in one and three dimension. The transformation was first calculated by hand to the fifth order, and later by a purpose built software program written in C++ to the eleventh order. The results of the transformation were finally generalized for any order by mathematical induction, and hence the sum of the expansion to all orders in the transformation was made possible. The coefficients of the hybridization terms calculated from this summation were later used to modify the original transformation, for eliminating the entire hybridization process from the model. The implications of these results to the PAM were discussed at the end.
List of Publications

The results of the research related to this work are, or soon to be, published in various journals under two major subjects.

The first subject is the exact transformation of the Anderson model or related models, described in chapter 6 and 7. The results are published in the following papers:


In addition to these papers, we have two manuscripts to be submitted in the near future.

The second subject is the high temperature superconductivity, to which our results have new revelations to the spin coupling between Copper and Oxygen holes, and the charge interaction between two on-site Oxygen holes. The papers published in relation to this subject are:


In this area, the future possibilities are endless, we have two manuscripts in preparation to summarize the results of chapter 4 applied to high temperature superconductors. In addition, the results for the three band Hubbard model have not yet been submitted. These results will most likely generate several more publications.
Contents

1 Introduction to Condensed Matter Physics 1
   1.1 A Brief History of Condensed Matter Physics ....... 1
   1.2 Four Fundamental Models of condensed matter physics .... 8
      1.2.1 The Hubbard Model ......... 9
      1.2.2 The $t - J$ Model ........ 12
      1.2.3 The Kondo Lattice Model .......... 13
      1.2.4 The Periodic Anderson Model .......... 15
   1.3 Recent Developments .................. 18

2 Introduction to Perturbation Theory 20
   2.1 Rayleigh-Schrödinger perturbation theory .......... 23
   2.2 Perturbative expansion of the small $U$ limit ...... 25
      2.2.1 First and Second order perturbations .......... 30

3 Perturbation Theory of the Single Site Limit 33
   3.1 Four electron states of a single atom .......... 34
      3.1.1 One-electron states .......... 34
      3.1.2 Two-electron states .......... 38
      3.1.3 Three-electron states .......... 47
   3.2 Perturbative expansion of the weak hopping limit .... 51
      3.2.1 First Order perturbation .......... 52
      3.2.2 Second Order Perturbation .......... 53
   3.3 Thermodynamic properties of the single site model .... 59
      3.3.1 Properties of canonical ensembles ...... 60
      3.3.2 Properties of grand canonical ensembles ...... 66
   3.4 Remarks ................................ 69
4 Perturbation Theory of the Atomic Limit

4.1 The Zhang-Rice Singlet ........................................... 73
4.2 Canonical Perturbation Theory ..................................... 83
  4.2.1 The coefficient of inter-subband hopping ....................... 88
  4.2.2 The inter-subband hopping .................................. 92
  4.2.3 Some interesting results .................................... 97
4.3 Thermodynamic properties ........................................ 102
  4.3.1 Haldane's perturbation expansion in single impurity model ... 103
  4.3.2 Magnetic susceptibility of the single and lattice impurity models 107

5 Introduction to Canonical Transformation .......................... 114

5.1 General Theory of Transformation .................................. 115
5.2 Renowned Applications to Condensed Matter Systems ............. 120
  5.2.1 Superconductivity ........................................... 121
  5.2.2 The t – J Model ........................................... 131
  5.2.3 The Schrieffer Wolff Transformation .................................. 135

6 One Dimensional Canonical Transformation .......................... 142

6.1 Canonical Transformation ........................................... 143
6.2 Proof by Induction ................................................ 150
6.3 Evaluating the Coefficients ........................................ 151
6.4 The Transformed Hamiltonian ....................................... 154
  6.4.1 Symmetric Case ............................................. 156
  6.4.2 Asymmetric Case ............................................ 164
6.5 Remarks ........................................................... 166

7 The Three Dimensional Periodic Anderson Model .................... 167

7.1 Canonical Transformation ........................................... 169
  7.1.1 Non-crossing hopping limit .................................. 170
  7.1.2 Recursive on-site interaction ................................ 176
7.2 Spin interaction ..................................................... 178
7.3 Elimination of higher order hybridization terms .................... 188
7.4 Complete elimination of hybridization terms ....................... 206

Conclusion ............................................................ 210

Appendices

x
CHAPTER 1

Introduction to Condensed Matter Physics

The nature of condensed matter has fascinated scientists since the early 1950s. Even now there are still plenty of open questions that are waiting to be answered. One typical example is certainly high temperature superconductivity which has attracted a great deal of attention from the condensed matter community. The difficulty in this field is the nature of the interaction involved. In most other fields of science one can tackle a problem from the ‘divide and conquer’ principle, dividing a large problem into smaller basic building blocks and then studying the problem from the understanding of these basic blocks. Collective phenomena in condensed matters however do not exist in a single particle. It is about the collective behaviour of a large number of particles in an appropriate condition and therefore very difficult to comprehend from the point of view of elementary particles. This explains why this field has started about a century ago, but only became the main stream of research in the last 50 years.

1.1 A Brief History of Condensed Matter Physics

A lot of progresses and developments in the field of condensed matter physics occurred in the last century. In the early 1900s, scientists were still puzzling about the abnormal behaviour of some metals and alloys. Superconductivity is one of the famous examples. Kamerlingh Onnes [73] discovered superconductivity in 1911 when he found that the electrical resistivity of mercury disappeared suddenly as the temperature dropped below 4.2 K\(^\dagger\). The rapid decrease of resistivity was half

\(^\dagger\)It was in fact Gilles Holst, a highly qualified physicist, who first discovered superconductivity when measuring the resistivity of mercury at 4.2K. This is quite clear from reference [71], but why Kamerlingh Onnes did not mention Gilles Holst in the corresponding publication and Gilles Holst
expected at the time given the vague theories of electrical resistivity, but a sudden disappearance of resistance was a total surprise.

With this entirely new state of matter, Onnes started to explore this new area with a few ideas. He turned a superconducting wire into a magnet and conducted some ingenious experiments with it. In a nutshell, he managed to make current flow in a superconductive coil for several hours without a battery, suspend a lead sphere freely in a magnetic field generated by persistent current in a superconducting ring, and make a small bar magnet hover above a superconductive lead dish. However, Onnes’ dream of using a superconductive coil to produce a very high magnetic field was dashed by the relatively low magnetic-field bearing capacity of pure metal superconductors. The search for new superconductors with higher magnetic flux holding capacity began.

Meanwhile, the race to understand and explain the new phenomenon caused by the superconducting material had also started. Meissner and Ochsenfeld [67] in 1933 found that instead of retaining magnetic flux, metals spontaneously repel any magnetic flux within it when they become superconductive. This disappearance of magnetic induction, together with the disappearance of resistivity violate the very basic assumptions of Maxwell’s theory, rendering Maxwell’s electrodynamic equation useless. An entirely new approach was needed to understand the new electrodynamic phenomenon. This was undertaken and completed by Fritz and Heinz London [63] within two years after the discovery of the Meissner effect. The central feature of London’s electrodynamics is the symmetry between the superconductive current and the steady magnetic field. In Maxwell’s electrodynamics a steady current is associated with a steady magnetic field, but a steady magnetic field does not result in a current. Instead a time varying magnetic field produces current. This asymmetrical relationship between a steady magnetic field and a steady current is the key difference between Maxwell’s theory and London’s proposal. In superconductors the relation between steady current and steady magnetic field is symmetrical in the sense that a steady magnetic field generates a steady current and vice versa.

At the same time progress had been made towards the understanding of the “magnetic threshold” and the technologies to push this threshold up for high field superconductors. The penetration of alloy by magnetic flux was found to be a gradual process, extending over a wide range of fields, as opposed to metal in the Meissner effect. Also, the magnetic threshold increases if the dimension of the superconducting materials decreases to a size comparable to the penetration depth of magnetic field. This prediction was based upon London’s electrodynamic equation did not dispute with Kamerlingh Onnes about the injustice, still remains unknown.
and a basic thermodynamic principle, which served as a good testing ground for London’s theories.

Theoretical work on the theory of superconductivity was however still very slow. Many different theories had been proposed, new axioms designed to elucidate the phenomenon but all failed to fit the facts. As Felix Bloch put it, “Every theory of superconductivity can be proved wrong.” and this axiom turned out to be right for a long period of time. Contrary to the expectation of new revelations, the final theory turned out to be based upon small findings and results accumulated over the years. The fact that the entropy of a persistent current is zero, for example, suggested a small gap must be in the available energy states at the Fermi surface. The energy involved must also be very small, as metal becomes superconductive only at very low temperature.

The crucial idea that lead to the final theory was independently developed by Fröhlich [34] in 1950 and Bardeen [13] in 1955 which suggested that two electrons can be attracted to each other through the vibrations the electrons created as they propagate in a lattice of positive ions. For instance, the small variation in charge of a few positive ions caused by the propagation of an electron can be harnessed by a second electron passing by if it travels at the exact same speed but in opposite direction of the first electron. This is similar to a resonance effect and is better explained as coherence. It is this attraction between two negatively charged electrons that lead Bardeen, Cooper and Schrieffer in 1957 to their final theory of superconductivity [15]. A year later Bogoliubov [19] in U.S.S.R. arrived at the same result using a much simpler method, canonical transformation. At the same time, Abrikosov [2] also published a full theory of superconductivity in 1957 for alloys, predicting in detail the behaviour of an ‘ideal’ alloy superconductor. The family of materials which exhibit this kind of behaviour later became known as type II superconductors.

Another interesting discovery around the same time in history is the superfluidity of helium. It is quite inconceivable that thirty years should have gone past between the first liquefaction of helium and the discovery of superfluidity, when many experiments with liquid helium were made in laboratories all over the world in those thirty years. As in the case of superconductivity, it was again the enormity of the discovery that prevented physicists from embracing it. The rapid change of specific heat and the density maximum at the transition temperature of 2.2K was spotted by Kamerlingh Onnes and Leo Dana as early as 1924, but due to the large magnitude of changes observed, they attributed the finding to an erroneous measurement and did not dare publishing it. Six years after the death of Onnes, the
1. Introduction to Condensed Matter Physics

same issue was published by Willem Hendrik Keesom and Klaus Clusius [52] who suggested the effect was the result of the formation of 'liquid crystal' which however turned out to be wrong.

The discovery of superfluidity for the explanation of this anomaly was quite an unexpected one when it had been noticed in a number of laboratories that containers containing liquid helium which were not perfectly tight, leaked excessively below 2.2K. Sometimes an apparatus which appeared quite tight above 2.2K would begin to leak badly as soon as it was cooled, and became useless. The measurement of the internal friction of liquid helium soon suggested the possibility of superfluidity. This explained not only the high specific heat of liquid helium, but also the sharp increase of heat conductivity in helium thin film. It was not the heat conductivity of the film which was high but an actual flow of liquid along it. The heat transfer in this case is not of conductive nature, but due to film transfer up the walls of the container and subsequent evaporation at the warmer region.

London later used Einstein's condensation theory of Bose gases to calculate the specific heat of helium liquid which lead Lazlo Tisza [82] to conclude that liquid helium contains two parts, the normal helium liquid and the 'condensed' helium superfluid. This 'two-fluid model' explained beautifully all the superfluidity phenomenon of liquid helium, except the physical reason for the existence of two fluids. For this, Lev Davidovich Landau [58] received the Nobel Prize twenty years later on the explanation of superfluidity. According to Landau there exists just one fluid; 'superfluid' liquid helium. As the temperature of the liquid is raised from absolute zero, thermal energy is provided in the form of vibrational quanta, phonons, from the walls of the container. These quantized vibrations of the helium atoms travel through the liquid. They slow down the fluid as soon as they scatter the liquid helium near the walls of the container, which essentially increases their effective mass. It therefore looks as if there were two different kinds of fluids coexisting in the same container, one being viscous and massive while the other one superfluid.

The discovery of superconductivity and superfluidity has captured the focus of scientists for at least the past 50 years and the accounts mentioned above is only an abridged version of the full history from reference [68]. Apart from these two major discoveries at low temperatures, another area of condensed matter physics that attracted a great deal of interest in the 1950s is the resistive minimum in alloys. For instance, the resistance minimum observed in Gold as recorded in 'The Theory of Metals' by A.H. Wilson [90] was commented "the cause of the minimum is entirely obscure and constitutes a most striking departure from Mathiessens's rule, according to which the ideal and residual resistances are additive; some new
1.1. *A Brief History of Condensed Matter Physics*

physical principle seems to be involved." Gold, being metal, is very conductive at low temperatures when the thermo-vibration of atoms is substantially reduced. In other words, electrons should find less resistance travelling in a lattice of atoms with minimal vibration. As temperature increases, the atoms in metal start oscillating around their equilibrium position which in turn increases the chance of hitting a travelling free electron and thereby increases the electrical resistance.

The observation of this 'obscure' resistance minimum in Gold is however contradicting this general understanding. It was only found out later that this resistance minimum is not a natural property of metals, but the effect of foreign particles with local spin moment dissolved into the metal. The question was then pointed to where the particles could have come from. The answer lies in the Gold purifying process. To get pure Gold in the past, one would have to melt the Gold nugget, found from Gold mine, into a high temperature bearing container such as cast iron bucket. As the Gold sets, some Iron atoms from the container 'migrate' into the otherwise pure Gold metal. Iron as we know has strong local magnetic moment, which account for the strong local spin moment of the impurity particles.

Theoretical work on impurities in metals in the late 50s concentrated on explaining the trends in the behaviour of impurity elements across the transition element series. 'Virtual bound states' was one of the most important concepts that emerged out of this work by J.Friedel [33]. These are states that are almost localized due to resonant scattering at the impurity site. In this state the conduction electrons spend a relatively large proportion of time in the region of the impurity, but it is not a bound state as the electron will eventually become free. This new concept allows some early interpretations of transition metal impurity effects in metal. However using virtual bound states alone, it is still not possible to explain many of the phenomena caused by impurities in metals, such as the resistance minimum in Gold.

P.W.Anderson [6] in 1961 looked into this problem from a different angle. He put forward a new idea which is now known as the Anderson Model, to account for all the underlying interactions between free electrons and electrons localized at the impurity site. The main features of the model are a short range interaction $U$ between the localized electrons, as well as a narrow resonance associated with the impurity states. Using the Hartree Fock approximation, he showed that the model can explain the ferromagnetic nature of some metals with impurities. His result was confirmed in the same year by P.A.Wolff [92] who looked at the problem from a different point of view. Wolff considered the problem as a scattering problem in which the conduction electrons scatter from the potential due to a single impurity
atom. The short range interaction $U$ in the model between the localized electrons is crucial to the creation of localized magnetic moments, even though it is not obvious from the Hamiltonian.

In 1963, Hubbard [48] also proposed that the correlation effects between electrons of narrow energy bands are important in determining the properties of transition and rare-earth metals. This effect, he emphasized, is more important than the corresponding effects in the conduction band. Since the theory for the correlation effect of free electron gas is different to the corresponding theory for electrons in narrow energy band, Hubbard suggested a model, later known as the Hubbard model, to account for the strong correlation between electrons of the same atom. The model also contains weak interaction between electrons of different atoms to reflect to some extent the behaviour of electrons in metal. Spins are associated with each electron in the model, to include the magnetic interaction into consideration. This model can in fact be transformed canonically to a $t - J$ model which shows explicitly the spin interaction between electrons. The nature of interaction between electrons in this model is simpler than that of the Anderson model, but the behaviour of the model is by no means less complicated.

The search for the theory of resistivity minimum was mainly focused on the s-d model, a model which contains a local magnetic moment associated with a spin $S$ coupled via an exchange interaction $J$ with the conduction electrons. Kondo [56] was the first in 1964 to show that this exchange interaction leads to singular scattering of the conduction electrons near the Fermi level and a $lnT$ contribution to the resistivity, using third order perturbative expansion with respect to the weak coupling constant $J$. Using this $lnT$ term with the phonon contribution to the resistivity, it is sufficient to explain the observed resistance minimum and hence the problem is solved.

The s-d model was later found by Schrieffer and Wolff [78] in 1966 to be a local moment limit of the Anderson model. The two models can in fact be related by a canonical transformation, the famous Schrieffer and Wolff transformation. The solution to the problem of the resistivity minimum for the s-d model can hence be mapped to the Anderson model in the local moment limit.

However, there is still one problem plaguing the proposed theory by Kondo as $lnT$ terms diverge when $T$ goes to zero. It is obvious that Kondo's perturbation calculations could not be valid at low temperatures. A more comprehensive theory was needed to explain the low temperature behaviour of systems giving resistance minima. The search for such a theory became known as the 'Kondo problem' and attracted a lot of theoretical interest in the late 60s and early 70s.
One of the approaches to the problem in the late 60s was introduced by Anderson who used the concept of scaling to remove the high energy spectrum. Anderson et al [7] showed that if the higher order excitations were eliminated perturbatively to give an effective model valid on a lower energy scale, the effective coupling between the local moment and the conduction electrons increased. This perturbational approach broke down when the coupling became so large that it could not be carried out down to the lowest energy scales at low temperature.

Thanks to the development of computers in the 1970s, numerical solutions became possible with problems that could not be solved analytically. Using computers, K.G. Wilson [91] in 1974 devised an algorithm, using ‘numerical renormalization group’, to confirm the observed power law behaviour in Kondo or s-d models. The method is non-perturbative and the idea is borrowed from field theory and scaling in critical phase transition. For this, he received a Nobel prize in 1982. The same technique was also applied to the Anderson model by H.R. Krishna-murthy, J.W.Wilkins and K.G.Wilson [57], in 1979 to resolve the remaining questions of single-impurity Anderson model. Only a year later, the exact solution to the Kondo model was found by N.Andrei [9, 10] using the Bethe ansatz, a hypothesis first used by H.A. Bethe in 1931 to solve the one dimensional Heisenberg model. This provides analytic results for the high and low temperature behaviour of the model which agrees with Wilson’s calculations. The exact solution for the Anderson model was also unravelled by P.B Wiegmann [88, 89] in 1982, using the same method.

These events marked the beginning to the end of the single impurity weakly interacting models in which quasi-particle, perturbation, Fermi liquid theory and scaling theories were successful in explaining many of the physical phenomena. There might still be on-going researches for the N-fold spin-orbit degenerated models, but the main focus has been shifted towards high temperature superconductivity.

Since the discovery of superconductivity in metals by Kammerlingh-Onnes in 1911, a great deal of researches have been directed towards the developments of compounds with high superconductive transition temperatures $T_c$. The progress was slow at the beginning because the nature of the interactions responsible for this kind of superconductivity at high temperature is different to the low temperature counterpart. The breakthrough came in 1985 when J. Bednorz and A. Müller [51], who had been experimenting in vain with many insulating perovskites and oxides, replaced the Ni$^{3+}$ ions in the metallic perovskites compound LaNiO$_3$ with Cu$^{3+}$ ions. In doing so, they managed to push the transition temperature of 11K from the first sample to 35K with samples that had been optimized in relative composition and thermal treatment. Nowadays, thanks to the impressive progress in both the
1. Introduction to Condensed Matter Physics

Science and the technologies, the transition temperature $T_c$ of some superconductive compounds has already reached 135K. In spite of all these experimental successes, the theoretical understanding of the high temperature superconductivity is still far from satisfactory. The interactions between the carriers, the localized particles and the phonons become so strong that a lot of the weakly interacting single impurity model become useless. Finding a model that would incorporate all the different anomalies observed in high temperature superconductors is a formidable challenge.

1.2 Four Fundamental Models of condensed matter physics

There are many anomalous behaviours of the high Tc superconductors that are far from understood. One type of these superconductors is the so-called 'heavy fermions' which has a very large specific heat coefficient and corresponds to a large effective mass. These compounds show diverse forms of low temperature behaviour; some appear to be unconventional superconductors, some order magnetically, some do both and others seem to remain paramagnetic. This anomalous behaviour is believed to be due to the strong on-site Coulomb interaction for the impurities, together with the 'mixed valence' or 'intermediate valence' state in which different valences of an ion co-exist in the compound. In this respect, a number of models have been put forward to explain this behaviour. One of them is the 'Kondo lattice' which puts emphasis on the spin interactions between the carrier and the localized particles. In forming a singlet ground state in this model, the conduction electrons must screen the impurity moment but the screening would span a distance much large than the interatomic distance. In low alloy concentration, this would not be a problem and the result would be similar to that of the single impurity model. However, for models with highly concentrated impurities, the conduction electron screening clouds overlap and the interimpurity interactions become significant. The physics of the concentrated system must therefore be quite different to that of the dilute alloys.

A similar model to this is the Periodic Anderson Model which includes Coulomb charge interactions between localized electrons and virtual resonance band between conduction and localized electrons. Many believe this model describes in a better way what is happening inside the superconductive compound, but the complexity of the model forbids a better understanding of its properties. Another famous model is the Hubbard model. The reason for its popularity is its simplicity. It contains only two terms, one responsible for the kinetic energy of the electron and another
for the Coulomb repulsion between electrons at the same site. Despite its simplicity, its solution in high dimensions still remains to be one of the most challenging tasks. In the following, we will describe some of the anomalous behaviours in different lattices, and introduce some of the models put forward to explain these behaviours.

### 1.2.1 The Hubbard Model

One of the minimal models used frequently to explain the electronic structure of CuO$_2$ planes in high temperature superconductor cuprates is the Hubbard Model. It is believed that the Hubbard model contains the essential physics of doped cuprates, whereas the three-band Hubbard model provides a more realistic description. In fact the Hubbard Model appears in many different forms, but the original one came first from Gutzwiller [39] in 1963 and later from Hubbard [48] himself, in a bid to describe in a simplified way the effect of correlations for d-orbital electrons in transition metals. In Hubbard’s original paper, he showed by means of an example how to model the effect of electron spin in an otherwise non-interacting model. He first associated the spins with the moving d-orbital electrons. According to Hund’s first rule for atoms, the intra-atomic interactions are of such a nature as to bring the electron spins of the atom to point at the same direction in half filling. Suppose that the electrons have their spins quantized in what will be called the up and down directions and that at some instant a given atom has its total spin in the up direction. Then the intra-atomic interactions are, according to Hund’s rule, of such a nature that this atom tends to attract electrons with spin up and repel those with spin down. In this way the property of an atom on having total spin at some instant tends to be self-perpetuating. If these intra-atomic forces are strong enough to produce appreciable correlations, then it follows that the state of total spin up on an atom may persist for a period long compared with the $d$-electron hopping time. In these circumstances one can think of the spin as being associated with the atom rather than with the electrons and the possibility of an atomic or Heisenberg model emerges [48]. He proposed what is now known as the Hubbard model in its simplest form:

$$H_{\text{Hubbard}} = \sum_{i,j,\sigma} t_{ij} c_{i,\sigma}^{\dagger} c_{j,\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow},$$

(1.1)

This Hamiltonian contains two contributions, a hopping integral $t_{ij}$ which is usually taken to be translationally invariant and acting between nearest-neighbours, i.e. $t_{ij} = -t$ for $i, j \in$ nearest-neighbours and zero otherwise, and an on-site term of strength $U$ representing the effective screened Coulomb interaction. The important component of this interaction term is $U \sum_i n_{i,\uparrow} n_{i,\downarrow}$, which gives an energy penalty.
1. Introduction to Condensed Matter Physics

(in the $U > 0$ case) for each double-occupied site. A great deal of effort has been devoted to the solution of this model, but exact results are still confined to the one-dimensional case.

The Hubbard model is believed to be capable of describing the main collective features of the itinerant magnetism and metal-insulator (Mott) transition, since when $U = 0$, $H$ reduces to a non-interacting system with free electrons, while when $t = 0$, the system is insulating at half-filling with one electron per site. In the subject of high temperature superconductors, it is believed that the correlation between the magnetic behaviour and the Mott transition near half-filling is crucial, as most of the ceramic superconductors are good Mott insulators and the CuO$_2$ plane in the superconductors exhibit strong antiferromagnetic correlations, when the material is close to superconductive.

The Hubbard model can also be mapped by canonical transformation onto the so called $t - J$ model, which will be discussed in the next section, to show the spin interaction inherited in the simple form of Coulomb interaction. The result of the transformation, after excluding the doubly occupied states at half-filling, reduces to the Heisenberg model for the antiferromagnetic insulators with $s = 1/2$.

However, it has been known for some time that the simple Hubbard model is not sufficient to describe electric and magnetic properties of oxides. In fact, one has to ignore the oxygens and the doping in the high temperature superconductors in order to fit the one-band Hubbard Model. The more realistic model must therefore contain the Cu$^{2+}$ orbitals $3d_{x^2-y^2}$ and the O$^-$ orbital $2p_x, 2p_y$ which gives rise to the three-band Hubbard Model. The Hamiltonian of the three-band model is then:

$$H_{3\text{band}} = \epsilon_d \sum_i n_{i}^d + \epsilon_p \sum_{j} n_{j}^p - t_{pd} \sum_{(ij)\sigma} (p_{j,\sigma}^\dagger d_{i,\sigma} + \text{H.C.}) + U_d \sum_i n_{i,d}^\uparrow n_{i,d}^\downarrow - t_{pp} \sum_{(jj')\sigma} (p_{j,\sigma}^\dagger p_{j',\sigma} + \text{H.C.}) + U_p \sum_{j} n_{j,p}^\uparrow n_{j,p}^\downarrow + U_{pd} \sum_{(ij)} n_{i,d}^\uparrow n_{j,p}^\downarrow.$$  (1.2)

The fermion operators $p_{j,\sigma}^\dagger (p_{j,\sigma})$ create (annihilate) O$^-$ holes $(2p_x, 2p_y)$ at site $R_j$. The fermion operators $d_{i,\sigma}^\dagger (d_{i,\sigma})$ create (annihilate) Cu$^{2+}$ $(3d_{x^2-y^2})$ holes at site $R_i$. $\langle i, j \rangle$ denotes nearest-neighbour pair. The phase factors are included in the summation definition. $n_{i,\sigma}^d = d_{i,\sigma}^\dagger d_{i,\sigma}, n_{j,\sigma}^p = p_{j,\sigma}^\dagger p_{j,\sigma}, n_{i,d}^d = n_{i,d}^\uparrow + n_{i,d}^\downarrow, n_{j,p}^p = n_{j,p}^\uparrow + n_{j,p}^\downarrow$, for $\sigma = \uparrow$ or $\downarrow$. The first added hole will go to the d-orbital of Cu$^{2+}$ ion and this is the case for undoped material such as La$_2$CuO$_4$. The next added hole will go to p-orbital of oxygen if $U_d \gg \Delta = \epsilon_p - \epsilon_d > 0$. Since $\Delta + 2U_{pd} < U_d$ doping with Sr$^{2+}$ will introduce a hole in O$^-$ $(2p)$ orbital. The overlap of the orbitals will result in delocalization of oxygen holes and a band is hence formed. However, a strong Coulomb repulsion localizes Cu$^{3d}$ holes and retains Cu$^{2+}$, but mobile holes are mostly on oxygen. Hopping induces superexchange interaction between Cu$^{2+}$ spins.
1.2. *Fundamental Models of condensed matter physics*

In this case, the immobile holes on Copper are responsible for antiferromagnetism, while superconductivity is due to the mobile oxygen holes. Such a simplified picture is supported by electron spectroscopy \[70\].

At half-filling with a copper hole per unit cell, the three band model reduces to the Heisenberg antiferromagnet \[97\], with the spin coupling \( J \) between carriers and localized holes given by:

\[
J = \frac{4t^4}{(\Delta + U_{pd})^2} \left( \frac{1}{U_d} + \frac{2}{U_p + 2\Delta} \right), \tag{1.3}
\]

The numerical estimate of \( J \) in this case is in good agreement with the value extracted from magnetic measurements.

However, intensive numerical works on the Hubbard model using exact diagonalization and Quantum Monte Carlo simulations have yet failed to find evidence at \( U > 0 \) for the spatial growth of pairing-correlations or evidence for a finite superfluid density, in the range of temperatures and cluster sizes that have been investigated \[29\]. In other words, there seems to be no clear indication of superconductivity in the realistic regime of parameters for either one-band or three-band Hubbard model.

Another variety of the Hubbard model is what is called the extended Hubbard model \[65\], in which an extra term is added to the Hamiltonian

\[
H_{eH} = H_{Hubbard} + \frac{1}{2} \sum_{i \neq j, \sigma, \sigma'} n_{i, \sigma} n_{j, \sigma'} \frac{e^2}{R_{ij} - R_j}, \tag{1.4}
\]

This extra Coulomb interaction term between electrons at different sites and different spins provides charge fluctuations into the original Hubbard model which would otherwise only contain spin fluctuations. Both the original Hubbard model and the extended Hubbard model allow for the separation of charge and spin degrees of freedom, which is exact in Random Phase Approximation (RPA) but only approximate in vertex corrections. Details of these can be found in standard text books such as \[65\]. The idea of spin and charge separation has been proposed as the normal conductive state of high temperature superconductors \[8\], in which an electron is a composite object consisting of a charged boson called holon, and a spinon, carrying a spin label. This idea is substantiated by the Bethe Ansatz solution and is used in the resonating valence bond theory. The same approach can also be applied to the \( t - J \) model using the slave boson representation.
1.2.2 The $t - J$ Model

Originated from the idea of Anderson [8] in 1987 that the high temperature superconductors can be modelled on the basis of the strongly correlated Hubbard model, the Hubbard model has been intensively investigated at the atomic limit $U > t_{ij}$. Perturbation and canonical transformation are widely used to show the antiferromagnetic ground state at half-filling. On the one hand, the electron hopping can be treated as a perturbation which at the second order yields an AF coupling due to the virtual hopping of electrons. On the other hand, one can perform a canonical transformation in such a way that the new quasiparticles have the property that although they do hop around, their hopping does not change the effective number of doubly occupied sites. The equivalence of these two different methods was demonstrated by Anderson [5] in 1959 in terms of the so-called kinetic exchange interaction.

For sufficiently large $U$, it is well understood that the bare energy band splits into two subbands. The lower subband is full while the upper subband is empty at half-filling, or one electron per atom. The virtual electron hopping between the upper and the lower subbands is the main cause of the antiferromagnetic coupling, which in real space corresponds to virtual electron hopping between a pair of singly occupied nearest-neighbour sites. The spins of the two electrons on this pair of sites are anti-parallel. This implies that the cross subbands hopping is responsible for the antiferromagnetic coupling. With respect to this, one can use a canonical transformation to remove the the cross subbands hopping from the Hamiltonian which retains only the motions of electrons in either the lower or the upper subband. The result of the transformation shows a spin interaction among the electrons at different sites when double occupied sites are ignored in the calculation.

$$H = -t \sum_{(i,j),\sigma} (c^\dagger_{i,\sigma} c^\ast_{j,\sigma} + \text{H.C.}) + J \sum_{(i<j)} \mathbf{S}_i \cdot \mathbf{S}_j,$$

(1.5)

Details of this equation and the derivation are shown in section 5.2.2 of chapter 5 which essentially follows the work of Chao, Spalek and Oleś [22].

The interest of this model in the subject of high temperature superconductivity is the description of low energy excitations by an effective single-band model in the strong correlated limit $U_d \gg t_{pd}$. Zhang and Rice [98] have shown that the oxygen hole and spin Cu$^{2+}$ ion can form a complex cluster due to exchange interaction, which can be singlet or triplet. They showed that the dynamics of the singlet, commonly known as the Zhang-Rice singlet can be described by the effective $t - J$ model. This leads to the Hamiltonian of d-orbital holes moving in the 2D square lattice of copper spins in the subspace of no doubly occupied sites with effective
hopping $t$ and spin coupling $J$. The theory is however based only upon nearest neighbour hopping when next nearest neighbour hopping could indeed be quite significant in some of the high temperature superconductors. The question of which model is the most suitable for the electronic structure for cuprate high temperature superconductors is still yet to be settled.

1.2.3 The Kondo Lattice Model

The Kondo lattice is one of the canonical models used to study strongly correlated electron systems, and has been the subject of intensive study. The Kondo lattice describes the interaction between a conduction band, containing Bloch-like delocalized electrons, and a lattice of localized magnetic moments. Its importance is due to its relevance to several broad classes of real materials, which includes heavy fermions and high temperature superconductors.

In the simplest case of the Kondo model, the conduction electrons hop between nearest-neighbour sites $(ij)$ only, and the Hamiltonian for the conventional Kondo lattice is written as

$$H_{KL} = -t \sum_{(ij)} \sum_{\sigma} c_{i\sigma}^\dagger c_{j\sigma} + J \sum_{j} S_{ej} \cdot S_{j}, \quad (1.6)$$

for a spin-isotropic interaction. $S_{ej}$ is the spin operator of the $c$-electrons and $S_{j}$ the spin operator of the localized electrons, both at site $j$ of the lattice. $t$ is defined as the hopping matrix element of the $c$-electrons between nearest neighbours, in the Hubbard model.

The Kondo lattice model contains two parameters. The first is the coupling $J$ (or the dimensionless parameter $J/t$). For large values of $J$ the relevant sign of $J$ is negative, and this is called a ferromagnetic coupling since it favours an alignment of the conduction electron spin with the spin of the localized $f$-electron. For small values of $|J|/t$ the relevant sign of $J$ is positive, and this is called an antiferromagnetic coupling since it favours an opposite alignment of the conduction electron spin with the spin of the localized $f$-electron. The second parameter in the Kondo lattice is the number of conduction electrons. This is measured by the filling $n = N_e/N$, where $N_e$ is the number of conduction electrons and $N$ is the number of lattice sites. $N$ coincides with the number of $f$-electrons.

One of the real materials that resemble the electronic structure of Kondo lattice is Manganese oxide perovskites. The key element of resemblance arises from the properties of the $3d$ shell electrons in Mn. The Kondo lattice describes the electrons in the $3d$ Mn shell as follows: A very strong Hund’s rule coupling forces the alignment of the spins of the localized $t_{2g}$ electrons, and these act as a localized spin $3/2$. 
The $t_{2g}$ electrons in the manganese oxides form the localized band in the Kondo lattice model, and as noted above are approximated by spins $1/2$ in the simplest case. The Kondo lattice conduction band models the delocalized $e_g$ orbitals. These are coupled to the localized electrons by a Hund's rule coupling; as for the Hund's rule coupling between the $t_{2g}$ electrons, the coupling is strong and favours an alignment of the $e_g$ electron spin with that of the localized $t_{2g}$ spins. Thus the parameter regime of the Kondo lattice which is appropriate for the manganese oxides is one with a large ferromagnetic coupling, \textit{i.e.} $J < 0$, $|J|/t > 1$.

Another class of compounds which is very similar to the Kondo lattice model is the heavy fermion which has the characteristic feature of specific heat being about two or three orders of magnitude greater than that of normal metals. This enhancement may be accounted for by supposing that the quasiparticle effective mass is two or three orders of magnitude greater than the bare electron mass, and thus the heavy fermions. Heavy fermion materials exhibit a striking diversity of ground states, including magnetically ordered states ($\text{CeAl}_2$ and $\text{U}_2\text{Zn}_{17}$), novel (non-BCS) superconducting states ($\text{CeCu}_2\text{Si}_2$ and $\text{UBe}_{13}$), and ground states which are neither magnetically ordered nor superconducting ($\text{CeAl}_3$ and $\text{UA}_2$). A detailed discussion of the early heavy fermion compounds and their experimental properties may be found in the review by Stewart [79]. A review discussion of effective theoretical models for heavy fermion materials is contained in reference [61].

Along with the heavy fermion compounds, there are several related classes of compounds containing rare earth and actinide elements. One is the class of Kondo insulators, and includes $\text{CeNiSn}$, $\text{Ce}_3\text{Bi}_4\text{Pt}_3$, $\text{TmSe}$, and $\text{UNiSn}$. These are small gap semiconductors in which the gap, of only a few meV, arises from hybridization between the $f$-electrons in the rare earth and actinide ions, and a conduction band. The Kondo insulators show very different behaviour from normal band insulators which have a gap of at least a few tenths of an eV. The Kondo insulators are reviewed in references [4] and [32]. Another class is the low-carrier-density Kondo systems which include the trivalent cerium monopnictides $\text{CeX}$ ($X = \text{P, As, Sb, Bi}$), and Uranium and Ytterbium analogues $\text{USb}$, $\text{YbAs}$, among others. (See reference [80] for a brief review.) These systems have carriers (either conduction electrons or holes) whose densities are much less than those of the magnetic rare earth or actinide ions. The low-carrier-density Kondo systems show heavy fermion behaviour including the extreme case of insulating states, and exhibit very interesting and complex magnetic properties. These include ferromagnetic ordering in one plane, and a complex stacking of the ferromagnetic planes throughout the crystal.

The Kondo lattice model is relevant to all of the compounds mentioned above;
1.2. Four Fundamental Models of condensed matter physics

heavy fermion materials, Kondo insulators, and low-carrier-density Kondo systems. This is due to the $f$-electrons from the rare-earth and actinide elements, which remain essentially atomic-like in the compounds. The rare earth elements from Cerium to Thulium have atomic configurations $[\text{Xe}]4f^n5d^{10}6s^2$ with partially filled $4f$ shells $2 \leq n \leq 13$. There is a similar progression in the actinide series as the $5f$ shell is filled. The partially filled $f$-shells lead to a variety of magnetic states for these elements. In compounds involving these elements the $f$ orbitals remain strongly localized, and essentially retain their atomic character. The magnetic effects present in the isolated atoms persist in many rare earth and actinide compounds. Thus the sites containing rare earth or actinide atoms often possess a magnetic moment corresponding to the Hund’s rule maximization of the total $f$-electron spin. In the compounds discussed above the $f$-electrons interact with electrons in the conducting $d$ (or hybridized $s$-$d$) band. The conduction band filling depends on the particular compound structure, and varies from small fillings in the low-carrier-density Kondo systems to half-filling for the Kondo insulators. The interaction between the conducting electrons and the localized electrons in the $f$ orbitals is the basis of the relevance of the Kondo lattice model to these materials.

Despite all of the successes with the Kondo lattice model, one substantial weakness of the model is that the Kondo lattice model neglects electron-phonon coupling. More precisely, it ignores interactions between the electrons and the vibrations of the underlying lattice of ions, which has been clearly recognized recently as an important factor in high temperature superconductivity. For instance, the Raman phonons, heat conductivity, isotope shift, geometrical sensitivity and observation of polarons in cuprate high temperature superconductors indicate that there exists a strong coupling between phonons and electrons in the superconductive phase.

1.2.4 The Periodic Anderson Model

The spin interaction in the Kondo model can in fact be formulated as an indirect consequence of charge interaction between the conduction and localized electrons via the more fundamental periodic Anderson model (PAM). To introduce this model, it is best to consider a concrete example, which is taken from the review of Varma [84]. SmS is an ionic semiconductor at atmospheric pressure, and contains Sm$^{2+}$ and inert S$^{2-}$ ions. The $4f^6$ band level$^\dagger$ $\varepsilon_f$ lies just below the (unoccupied) $5d$ and $6s$

$^\dagger$The bands corresponding to the $4f$ orbitals are energetically narrow, or atomic-like, with a delta-function density of states $\rho(\varepsilon) \approx \delta(\varepsilon - \varepsilon_f)$, where $\varepsilon_f$ is the $f$ level. Note that these are not bands in the usual sense; band theory begins from a basis of non-interacting delocalized Bloch states (cf. section 1.1.1), and provides a poor description of the properties of partially-filled $4f$
Introduction to Condensed Matter Physics

conduction bands in Sm. If pressure is applied, the lower of the crystal-field split 5d bands of Sm broadens, due to the increased overlap of its Wannier states, and moves down in energy relative to \( \varepsilon_f \). Ultimately the 5d band crosses the level of the 4\( f^6 \) band. When this occurs, there are valence fluctuations on each Sm site, as f-electrons enter 5d conduction band states, and SmS becomes metallic. The metal-insulator transition is accompanied by a change in the colour and volume of SmS. The central lesson of this transition is that in order to understand the metallic state, it is necessary to understand excitations from the 4\( f \) localized band into the 5d conduction band. This situation of fluctuating valence is described by the periodic Anderson model, which considers interactions in which localized electrons may be excited to the conduction band, and vice versa, in interband interactions.

The periodic Anderson model is the formal extension to the lattice of the single-impurity Anderson model [6]. The general properties of the single-impurity model and their derivations are summarized in Appendix A, based upon the original paper [6] by P.W. Anderson in 1961. The model describes a band of conduction electrons together with localized f orbitals at each lattice site. The interaction between the localized and extended states is distinct from that in the Kondo lattice model, and describes excitations into and out of the localized f orbitals. On a lattice with \( N \) sites, the periodic Anderson Hamiltonian is given by

\[
H_{\text{PAM}} = \sum_{k,\sigma} \varepsilon(k) c_{k,\sigma}^\dagger c_{k,\sigma} + \varepsilon_f \sum_{j,\sigma} n_{j,\sigma}^f + U \sum_j n_{j,\uparrow}^f n_{j,\downarrow}^f + N^{-1/2} \sum_{k,j,\sigma} \left( V_k e^{i k j a} f_{j,\sigma} c_{k,\sigma} + \text{h.c.} \right),
\]

(1.7)

where the conduction electrons are written in terms of Bloch states with dispersion \( \varepsilon(k) \). \( \varepsilon_f \) is the level of the flat band of localized f orbitals. The hybridization \( V_k \) gives the amplitude for a localized f-electron to be excited to a conduction band Bloch state with crystal momentum \( k^\dagger \). An important element in \( H_{\text{PAM}} \) is Coulomb repulsion \( U \) between on-site f-electrons. This is by far the strongest interaction in real materials, \( U \sim 5 \text{ eV} \) [84], and large \( U \) will be assumed in the following.

The hybridization between the conduction band and the localized orbitals can shells, in which the f-electrons are localized and interact strongly with each other. For comparison with the energies of well-defined conduction bands, it is however convenient and conventional to consider a 4\( f \) 'band' for a given compound with a 4\( f^n \) nominal occupation on constituent rare earth atoms. The f level \( \varepsilon_f \) is then the energy for the process 4\( f^n \rightarrow 4 f^{n-1} \) of removing an f-electron. For example, in SmS \( \varepsilon_f \) is the energy for the process 4\( f^6 \rightarrow 4 f^5 \) [43, 84].

\(^\dagger\) According to standard band theory, \( V_k \) would be zero, since the states in different bands are orthogonal. However, as noted above, the f orbitals are strongly correlated, and do not constitute a band in the rigorous sense.
be an important element in rare earth and actinide compounds, as is clear from the example of SmS above. Hybridization in fact makes the Kondo lattice as a limiting case of the periodic Anderson model, in which the \( f \)-electrons are fixed at their lattice sites, and excitations from localized orbitals to the conduction band are forbidden. Schrieffer and Wolff [78] showed that the Kondo lattice is equivalent to the periodic Anderson model in the local moment regime. The local moment regime has the \( f \) level \( \varepsilon_f \) below the conduction electron chemical potential. In this case, for zero hybridization \( V_k = 0 \), the ground-state consists of one localized electron in each \( f \)-orbital, and a non-interacting Fermi sea of conduction electrons with chemical potential \( \varepsilon(k_F) \) at zero temperature. (\( k_F \) is the conduction electron Fermi momentum.) The local moment regime thus has \( \varepsilon_f - \varepsilon(k_F) < 0 \) and \( \varepsilon_f - \varepsilon(k_F) + U > 0 \). (Large \( U \) is assumed.) Details of the Schrieffer and Wolff transformation [78] is shown in section 5.2.3 of chapter 5. The Schrieffer and Wolff transformation gives a weak antiferromagnetic Kondo coupling for large \( U \) in the local moment regime [43]:

\[
J_{k,k'} = V_k^* V_{k'} \left\{ \frac{1}{U + \varepsilon_f - \varepsilon(k')} + \frac{1}{\varepsilon(k') - \varepsilon_f} \right\} \geq 0. \tag{1.8}
\]

The Kondo lattice Hamiltonian \( H_{KL} \) with a weak antiferromagnetic coupling (\( J > 0, J/t < 1 \)) thus describes rare earth and actinide compounds with weak valence fluctuations in the local moment regime.

The Anderson lattice model has also been proposed as a more realistic model for the high temperature superconductors [97]. In this proposal, the \( f_{j,\sigma}^\dagger \) creates holes in the \( d_{x^2-y^2} \) orbital on Cu site \( j \) whereas the \( c_{j,\sigma}^\dagger \) creates holes in a \( p_{\sigma(y)} \) orbital on O site \( j \). The hybridization now becomes the coupling or hopping of holes between the \( d \) and the \( p \) orbitals, whereas \( t \) is the hopping coefficient of holes between different \( p \) orbitals. The on-site repulsion \( U \) describes the Coulomb interaction between two \( d \) electrons at the same site of opposite spin. The important parameters of this model to the real materials are the Coulomb interaction \( U \) and the charge-transfer energy \( \Delta = \epsilon_f - \epsilon(k) \) which is equivalent to the energy cost of transferring a hole from the Cu site to the O site. When \( \Delta \gg U \), the \( 2p \) degrees of freedom can be projected out and the system becomes an effective one-band Hubbard model with transfer integrals of \( V^2/\Delta \). However when \( \Delta \ll U \), the conduction band still corresponds to the upper band but the valence band is part of the \( 2p \) orbital. The lower impurity band is lying even below the bottom of this valence band. When \( \Delta \approx U \), the upper and the lower bands and the \( p \) band are strongly mixed which drives the system into the mixed valence regime. Finally, if \( \Delta \) or \( U \) is of the order of the relevant bandwidth \( V^2/\Delta \), the system is well into the metallic regimes of either mixed-valence (for small \( \Delta \)) or Brinkman-Rice [20] (for small \( U \)).
1.3 Recent Developments

Extensive studies on the Hubbard model (which includes \( t - J \) model) and the Kondo lattice model in the last few decades has led scientists to a good understanding of the model in relation to the strongly correlated effects. In particular, the Hubbard model in one dimension was solved exactly in 1968 [62] and a great deal of numerical results using exact diagonalization and quantum Monte Carlo (QMC) simulation have been published for the two and three dimensional case in the last thirty years. Approximations were also made in the strong coupling limit when \( U/|t| \gg 1 \) to find the limiting behaviour of the model. The Gutzwiller variational approach [40] was one example of finding the ground state of the model, whereas the auxiliary field (or slave boson) approach [11] was another that had success in reproducing results from the Gutzwiller approximation and other mean-field theories. Both paramagnetic and ferromagnetic phases were predicted from the model with different fillings, but without an exact solution, the results are still controversial.

Opposite to the strong-coupling regime, the weak-coupling region has also been studied extensively. The ground-state energy in low dimensions can be obtained by diagrammatic perturbation techniques and self-consistent approximations. For intermediate values of the coupling constant, mean-field theory is the only one that is available. It provides a rich phase diagram at \( T = 0 \) with paramagnetic, ferromagnetic and antiferromagnetic ground state, depending on the filling of the band.

With the \( t - J \) model, it has been studied numerically in one dimension and is exactly solved at \( J/t = 2 \). In this case, no phase transition was found and no phase separation for \( 0 \geq n \geq 1 \). When \( J/t \leq 1 \), there is no indication of superconductivity, instead the magnetic correlations dominate. For \( J/t > 2 \) pairing correlations develop as a precursor of phase separation which occurs for \( J/t > 2 \) [76]. In two dimensions, the \( t - J \) model does not superconduct at small \( J/t \) and hole density [29]. At large \( J/t \) there is a phase separation or clustering of doped holes. In between there is a possibility of a superconductive state of d-wave symmetry.

Many numerical simulations exist in the literature for the Kondo lattice, using quantum Monte-Carlo and Renormalization group approaches, but rigorous results are still rare, due to the complicated interactions between electrons in the model. The phase diagram of the Kondo lattice in one dimension has been established by Honner and Gulács [44] rigorously which shows clearly the ferromagnetic and the paramagnetic phase. It also provides a non-perturbative explanation of the \( 2k_F \) correlations in the RKKY regime without the divergences that plague the perturbative approach. A similar result is also obtained from the density-matrix renormalization-group approach by McCulloch and Gulács [66] which shows an additional newly
1.3. Recent Developments

discovered intermediate ferromagnetic coupling in the region close to half-filling. This exciting discovery provides a hint of a new spin coupling interaction which has not been explained yet by any contemporary theory.

The Periodic Anderson model however remains to be one of the difficult model to tackle. It has a very wide parameter space which is capable of explaining many different interaction between electrons or holes in real materials. The bulk of this thesis is therefore about understanding the different interactions in the Periodic Anderson Model and searching for clues that could explain the anomalies of the model from numerical simulations. In particular, we found that some of the canonical perturbative expansions can be summed to infinite order and yield surprisingly simple results. This kind of summation happens not only to the main result when we consider the canonical transformation of Anderson model, but also to the canonical perturbative expansion of the same model in terms of energy subbands. The two approaches are very different in nature with different restrictions imposed on the model parameters. The same result that they share, seems to suggest that there is a more fundamental principle underneath these calculations that is causing the coincident.

In the next two chapters we will describe the periodic Anderson Model in two different perturbation limits. The first perturbation is considered based on the small $U$ limit. This limiting case does not seem to find many application in real materials, possibly due to the fact that the system is close to non-interacting. We will also consider the model under the weak hopping limit, $\epsilon(k) \to 0$, in the next chapter in which the PAM is solved with a single site. Perturbation is then performed on the kinetic energy of the Hamiltonian which shows that the system is antiferromagnetic in the second order perturbation. The atomic limit of PAM is followed in the next chapter that uses the canonical perturbation theory to find the interaction of electrons between different energy subbands. Diagrammatic techniques are developed and used to find the spin interaction between electrons to any order of the perturbation series. The result is however true only when the hybridization constant $V$ is small. The corresponding thermodynamic properties will also be discussed. This is followed by the introduction of canonical transformation in chapter four, on which our main results in chapter five and six are based.
CHAPTER 2

Introduction to Perturbation Theory

An exact solution for the periodic Anderson model has not yet been found. In the search for an exact solution, perturbation is an important tool with which we can investigate the behaviour of the model in certain extreme conditions, and from which one can postulate the solution that would not have otherwise been found from any methods currently used. The same technique had been applied to many different types of models, such as the Kondo Model [56] and the Hubbard Model [22] and the results have been promising. Although the theory is valid only when the perturbation is small, it does provide plenty of vital information about the model in different parameter space.

There are many different ways perturbation theory can be applied to a model. One famous example is undoubtedly the explanation of the resistance minimum in metal alloys by Kondo. He observed from experiments that the resistance minimum is a consequence of the interaction between the spins of localized and conduction electrons. In detail, the depth of the resistivity minimum relative to the resistivity at $T = 0K$ is roughly proportional to the impurity concentration. But its ratio with the resistivity at $T = 0K$ is independent of the same impurity concentration. This minimum, he concluded, must therefore be a single impurity effect, not due to the interaction between impurities. He calculated the conductivity of the s-d model perturbatively to second order of the spin coupling $J$ which yielded the same result as those for purely potential scattering. He then continued to extend the perturbation to the third order and found that the spin degeneracy of the impurity has a very significant effect on the low energy scattering of the conduction electrons. Detailed calculation [43] shows that the resistivity of the s-d model has a $-ln(kT/D)$ term which significantly reduces the overall resistivity at small $T$, causing a minimum out of a monotonously increasing $T^5$ contribution from the
phonon. This logarithmic term diverges however at temperature $T$ close to zero, which renders the perturbation invalid. The problem of finding a solution when $T \to 0$ is referred to as the 'Kondo problem'. The logarithmic divergence arising from the third order perturbation may in fact be removed if one would sum all the higher order logarithmic terms in the perturbation expansion. Abrikosov [3] in 1965 carried out this summation of the leading order divergent terms for the resistivity and found that the logarithmic terms were then in the denominator due to the summation of a geometric series. The difficulties arising from the $\ln T$ terms became more severe as there was a divergence at a finite temperature $T_K$, now known as the 'Kondo temperature'. This problem was finally solved using renormalization group theory predominantly pioneered by Wilson [91] in the early 1970s, based upon the scaling idea of critical phenomena in phase transitions introduced by Kadanoff [50].

There are many other examples of Green's function perturbation or self-energy perturbation to search for the physical properties underlying the model [39, 45, 46, 48]. The formulations are elaborated and the results very effective. We will, however, start with a more direct and simple way of finding out the behaviour of the periodic Anderson model under different small perturbative parameters. Let us first assume that the Hamiltonian has two Hermitian parts:

$$H = H_0 + H_1$$

(2.1)

Of these $H_0$ is regarded as the unperturbed part which can be solved exactly, and $H_1$ as the small perturbation term. In most cases, the contribution of $H_1$ to the solution is small and one can expand the solution using Taylor series around the solution of the unperturbed part, with respect to the small parameter $H_1$. The main idea of the perturbation theory is to find the expression for energies and wavefunctions of the stationary states in terms of the known energy value $E_n^0$, and the wavefunctions $\phi_n$, of the unperturbed system $H_0$.

If the part $H_1$ is indeed significant relative to the energy level of the unperturbed system, perturbation theory will break down as in the Kondo problem, and the solution would not reflect the actual behaviour of the model. Therefore, choosing the appropriate perturbed term $H_1$ from the Hamiltonian is crucial to the validity of the solution. In this respect, there are three different limits of the Anderson model to which perturbation can be applied, depending on the relative significance of the parameters.

I) The first limit is when the on-site Coulomb interaction of the localized electrons is weak, i.e. $U$ is small in Eq. (1.7). Double occupation of electrons at any localized site does not incur significant penalty to the system which
as a result causes the up and the down spin electrons to propagate as almost independent particles without mutual interactions. Yamada etc. [93, 94] had considered this case extensively using a polynomial approximation to replace the Gaussian function in the localized electron Green's function for the single impurity Anderson Model. This approximation simplifies the calculation enormously and provides a solution very close to the exact one. The result, however, cannot be generalized to the Periodic Anderson Model (PAM) without much considerations and alterations. Furthermore, the parameter $U$ for PAM is generally large and so perturbation on $1/U$ in fact becomes more realistic [12].

II) The second limit is when the hybridization between the conduction and the localized electrons is small, i.e. when $V \approx 0$ in Eq. (1.7). In this case, the probability of electrons tunnelling between the spherical and the d-orbital or the f-orbital is so small that electrons in these orbitals behave essentially as independent particles. Haldane [41] used this approach to find the partition function of the Hamiltonian for the single impurity Anderson model. His method is indeed suitable for models of a periodic nature, however this has never been explored until now. We found that based on his theory, the partition function of the Periodic Anderson Model is highly related to that of the single impurity model. From a totally different point of view, Jan Zaanen and A M Oleś [97] had extended the canonical perturbation theory used initially for the Hubbard Model to the Periodic Anderson Model for small $V$ limit within the lower order terms. Interesting results were found when we carried out the expansion using diagrams and summed the relevant diagrams to high orders. This approach has again never been attempted before. Both of these two methods are explained in detail in chapter 4.

III) The third limit lies in the region where the elements of the hopping matrix $t$ are small, leading to a model with slow moving particles. This idea was investigated by Bastide and Lacroix [12] who considered the Periodic Anderson Model with $U \rightarrow \infty$ for its dynamic interactions. We followed their approach in chapter 3 but considered the most general Periodic Anderson Model, without any restriction on $U$. The magnetic property of the model was evaluated from the difference in energy between electrons of similar and opposite spins. The complexity of the solution was far greater than the one considered in [12] as $U$ was arbitrary which increased the number of degrees of freedom by one. Some new results found from the calculations were simulated numerically to help visualize the behaviour of the non-linear results.
2.1. Rayleigh-Schrödinger perturbation theory

In the following, we will first introduce the Rayleigh-Schrödinger perturbation theory for physical systems with discrete energy levels that can be separated into two Hermitian parts as in Eq. (2.1). The solution for the first limit is discussed even though no realistic application is implied. The unperturbed system was solved and attempts to higher order perturbation were made afterwards. This is followed by the discussion of the third limit in chapter 3 which provides more insight into realistic physical models. The magnetic susceptibilities were derived from the formulation which provided some understanding of the model. The second limit is discussed in chapter 4 in which canonical perturbation theory is explained in more detail.

2.1 Rayleigh-Schrödinger perturbation theory

Assuming the unperturbed part of the Hamiltonian is non-degenerated, one can write

\[ H_0 \phi_n = E_n^0 \phi_n \]  \hspace{1cm} (2.2)

and let

\[ H_1 = \lambda W \]  \hspace{1cm} (2.3)

where \( \lambda \) is a small, dimensionless parameter. The task of finding the eigenvalues and eigenfunctions of the Hamiltonian operator becomes solving the equation

\[ (H_0 + \lambda W) \varphi = E \varphi \]  \hspace{1cm} (2.4)

If we change now from the coordinate representation to the energy representation and choose as our basis system, the set of eigenfunctions \( \phi_n \) of the operator of the unperturbed system, we get

\[ \varphi = \sum_n a_n \phi_n \]  \hspace{1cm} (2.5)

and Eq. (2.4) reduces to the infinite set of algebraic equations

\[ (E - E_m^0) a_m = \lambda \sum_n W_{mn} a_n \]  \hspace{1cm} (2.6)

where \( W_{mn} = \langle \phi_m | W | \phi_n \rangle \) are the matrix elements of the perturbed part of the Hamiltonian \( H_1 \).

To determine the corrections of the energy and wavefunctions of the stationary states using this new basis system, we can write

\[ E = E_0^1 + \lambda E_1^1 + \lambda^2 E_2^2 + \ldots \]  \hspace{1cm} (2.7)

\[ a_m = \delta_m^1 + \lambda a_m^1 + \lambda^2 a_m^2 + \ldots \]  \hspace{1cm} (2.8)
Substituting these series back to the Eq. (2.6), it is easy to write

\[
(E_i^0 - E_m^0 + \lambda E_i^1 + \lambda^2 E_i^2 + \ldots)(\delta_{ml} + \lambda a_m^1 + \ldots) \\
= \lambda \sum_n W_{mn}(\delta_{nl} + \lambda a_n^1 + \ldots)
\]

(2.9)

Putting \(m = l\) and comparing terms of the same order in \(\lambda\):

\[
E_i^1 = W_{ll},
\]

(2.10)

\[
E_i^2 + E_i^1 a_i^1 = \sum_n W_{ll} a_n^1,
\]

(2.11)

If \(m \neq l\):

\[
a_m^1(E_i^0 - E_m^0) = W_{mn},
\]

(2.12)

\[
E_i^1 a_m^1 + (E_i^0 - E_m^0)a_m^2 = \sum_n W_{mn} a_n^1,
\]

(2.13)

It follows from Eq. (2.11) that in first approximation the energy of the system is expressed by the formula

\[
E = E_i^0 + \lambda E_i^1 l = E_i^0 + \lambda W_{ll}
\]

(2.14)

The first-order correction to the energy is thus the average value of the perturbation operator \(H_i\) in the state corresponding to the zero\(^{th}\) order wavefunction \(\phi_i\). Using Eqs. (2.13), (2.3) and (2.5), we get for the first-order wavefunction of the state \(l\),

\[
\phi_l = \phi_i + \lambda a_i^1 \phi_i + \sum_{m \neq l} \frac{W_{mi}}{E_i^0 - E_m^0} \phi_m.
\]

(2.15)

The quantity \(\lambda a_i^1\) is determined from the normalization condition. The \(\phi_i\) are normalized. Neglecting terms of order \(\lambda^2\) or higher we get from the requirement that \(\phi_i\) be normalized:

\[
a_i^1 + a_i^1* = 0.
\]

(2.16)

Wavefunctions are determined apart from a phase factor, and from Eq. (2.16) it follows that \(a_i^1\) must be purely imaginary. We can therefore put \(a_i^1\) equal to zero. The first-order wavefunction is thus given by the equation.

\[
\varphi_l = \phi_i + \sum_{m \neq l} \frac{W_{ml}}{E_i^0 - E_m^0} \phi_m.
\]

(2.17)
2.2. Perturbative expansion of the small $U$ limit

Substituting the values of $a^1_n$ obtained from the first Eq. (2.13) into the second-order Eq. (2.11), the second-order correction to the energy is:

$$E_i^2 = \sum_{n\neq l} \frac{W_{ln} W_{nl}}{E_i^0 - E_n^0}. \quad (2.18)$$

The second-order expression for the energy when $\lambda = 1$ is thus

$$E_i = E_i^0 + W_{il} + \sum_{n\neq l} \frac{W_{ln} W_{nl}}{E_i^0 - E_n^0}. \quad (2.19)$$

It follows from Eq. (2.19) that the second-order correction to the energy of the ground state is always negative as the excited state is always at a higher energy than the ground state.

The perturbation theory method discussed here is valid only if the series of successive approximations converges. A necessary condition for this is that each correction is small compared to the preceding one. We can therefore write the condition for the applicability of perturbation theory in the form

$$|H_{lm}| = |W_{lm}| \ll |E_i^0 - E_m^0|, \quad \text{for } m \neq l \quad (2.20)$$

This condition for the applicability of the perturbation theory method leads to the requirement that the off-diagonal matrix elements of the perturbation operator $H_1$ are small compared to the absolute magnitude of the difference of the corresponding values of the unperturbed energy. In other words, the distance between the given energy level and all other energy levels of the unperturbed system must be large compared to the change in the energy caused by the perturbation. The applicability of Eqs. (2.17) and (2.19) is not violated even if some of the states $m$ with an energy $E_m^0$ satisfying Eq. (2.20) are degenerated [30]. These formulae can also be extended to the case where some of the states $m$ belong to a continuous spectrum; in that case the sums in Eqs. (2.17) and (2.19) must be replaced by integrals for those continuous states [30].

2.2 Perturbative expansion of the small $U$ limit

With the periodic Anderson model, there are three different possibilities to perturb the Hamiltonian for three different limits as mentioned at the beginning of this chapter:

- When $U$ is small, let $H_1 = U \sum_i n_i^ \dagger n_i^\prime$ and perturb the solution of the Hamiltonian with $U = 0$,
• When $V$ is small, let $H_1 = V \sum_{i\sigma} f_{i\sigma}^\dagger c_{i\sigma} + \text{H.C.}$ and perturb the solution of the Hamiltonian with $V = 0$.

• When $t$ is small, let $H_1 = t \sum_i c_{i+1}^\dagger c_i + \text{H.C.}$ and perturb the solution of the Hamiltonian with $t = 0$.

In this section, the ground and excited states of the electrons will be pursued, together with their corresponding energies in these three different limits. In general, if the ground state energies of up spin and down spin electrons are degenerated in the unperturbed system, higher order perturbation will be required to refine the energy difference between the two states. The difference of the energy eigenvalues between spin singlet and spin triplet configurations of the electron wavefunction is the main key to the understanding of the magnetic property of the model.

The first limit is discussed in this chapter, and the other two limits will span the next two chapters. We first discuss the case when $U$ is small, using the Hamiltonian (1.7) written in the momentum space\(^1\). The single impurity version of the unperturbed Hamiltonian $H_0$ was solved exactly in 1961 by Anderson [6] and Fano [31]. Anderson applied the model to solid-state physics, while Fano used it in atomic spectra. The details of the solution can be found in standard textbooks, such as [65], [43]. The solution shows that the model is merely a description of a localized scattering resonance. The continuum particles come to the impurity, spend some time in the resonant state, and then depart to another continuum state.

Here the approach to the periodic model is however slightly different. The asymmetric character of the localized and conduction electron in an otherwise perfectly symmetric Hamiltonian is first removed. The fact that the localized electrons are immobile breaks the symmetry of the localized and conduction electron in the Hamiltonian and causes the Fourier transform of the Hamiltonian to become unnecessary complicated. This can be altered by adding an extra kinetic term to the localized electrons [16, 36],

$$-t_f \sum_{i\sigma} (f_{i\sigma}^\dagger f_{i+1\sigma} + \text{H.C.}) \quad (2.21)$$

and allowing $t_f$ to approach zero later in the formulation.

\(^1\)This is the only section throughout the whole thesis that the PAM is investigated under the momentum space. The formulation in this section is therefore slightly different to the rest of the thesis.
Perturbative expansion of the small U limit

The unperturbed part of the Hamiltonian in momentum space is therefore:

\[ H_0 = -2t \int \frac{dk}{2\pi} \langle \cos k \rangle c_{k\sigma}^\dagger c_{k\sigma} - 2t_f \int \frac{dk}{2\pi} \langle \cos k \rangle f_{k\sigma}^\dagger f_{k\sigma} + \epsilon_f \int \frac{dk}{2}\pi f_{k\sigma}^\dagger f_{k\sigma} \]

\[ + V \int \frac{dk}{2\pi} (f_{k\sigma} c_{k\sigma} + c_{k\sigma}^\dagger f_{k\sigma}) \]  

(2.22)

The symmetric nature of this equation allows Eq. (2.22) to be written in a matrix form

\[ H_0 = \int \frac{dk}{2\pi} \begin{pmatrix} c_{k\sigma}^\dagger f_{k\sigma} \end{pmatrix} M \begin{pmatrix} c_{k\sigma} \\ f_{k\sigma} \end{pmatrix} \]  

(2.23)

in which \( M \) is given by

\[ M = \begin{pmatrix} -2t \cos(k) & V \\ V & -2t_f \cos(k) + \epsilon_f \end{pmatrix} \]  

(2.24)

One can easily diagonalize this matrix with \( t_f \) set to zero, and solve for the eigenvalues

\[ E_{\pm}(k) = \frac{(\epsilon_f - 2t \cos(k)) \pm \sqrt{(\epsilon_f + 2t \cos(k))^2 + 4V^2}}{2} \]  

(2.25)

with the corresponding eigenstates

\[ \begin{pmatrix} \alpha_\pm(k) \\ \beta_\pm(k) \end{pmatrix} = \begin{pmatrix} V \\ \sqrt{(E_{\pm}(k) + 2t \cos(k))^2 + V^2} \\ \sqrt{(E_{\pm}(k) + 2t \cos(k))^2 + V^2} \end{pmatrix} \]  

(2.26)

in which \( \alpha_\pm \) and \( \beta_\pm \) are the coefficients of the \( c \) and \( f \)-electrons in high eigen-energy state \(+\) and low eigen-energy state \(-\). The two coefficients are also assumed normalized, \( \alpha_\pm^2 + \beta_\pm^2 = 1 \).

The Hamiltonian can hence be reformulated into:

\[ H_0 = \sum_{\mu=\pm,\sigma} \int \frac{dk}{2\pi} \Phi_{\mu\sigma}^\dagger(k) E_\mu \Phi_{\mu\sigma}(k) \]  

(2.27)

where the fermion operators \( \Phi_{\mu\sigma} \) are transformed from the original conduction and localized electron operators as,

\[ \begin{pmatrix} \Phi_+(k) \\ \Phi_-(\sigma) \end{pmatrix} = \begin{pmatrix} \alpha_+(k) & \beta_+(k) \\ \alpha_-(k) & \beta_-(k) \end{pmatrix} \begin{pmatrix} c_{\sigma}(k) \\ f_{\sigma}(k) \end{pmatrix} \]  

(2.28)

This equation shows that the conduction and the localized electrons have form two combined new states which can be considered as two new quasiparticles \( \Phi_{\pm\sigma} \), one of them sitting at a higher energy level than the other. The dispersion relationship of both energy levels is depicted in figures 2.3 and 2.4. When \( \epsilon_f \) is very...
negative, the lowest energy level of the system has the total energy proportional to $\epsilon_f$. This indicates the system is mainly populated by the localized f-electrons, which is consistent with the dominance of the localized electron in the eigenstates of the diagonalized Hamiltonian in figure 2.2. The relative populations shown in figures 2.1 and 2.2 are the magnitude square of the coefficient $\alpha_+ \beta_\pm$ in the corresponding eigenstate of the Hamiltonian $(\alpha_+^2, \beta_\pm^2)$. The offset of $\epsilon_f$ from zero at the half filled case is due to kinetic energy of the conduction energy $2t \cos(k)$ which changes the energy gap between the conduction and the localized electrons.

When $\epsilon_f$ is very large at the other end of the scale, the energy level of the lower energy state approaches zero and the system becomes dominated by the conduction electrons, as indicated on the right hand side of figure 2.4. Figure 2.1 shows the relative populations of the two electrons in this higher energy eigenstate. This figure is in fact identical to figure 2.2, except for the labelling, which is the result of the symmetry between c and f electrons in the model.

In between these two extreme cases, localized electrons and conduction electrons interchange via the hybridization term, to attain the minimum energy of the system.
2.2. Perturbative expansion of the small $U$ limit

The second effect introduced by the hybridization is 'spin-flip', by which the spin of the localized electrons changes from one state to another. For example, a local down spin electron can jump into a conduction state with spin down and then drift away. Some time later, a different conduction electron with spin up can come and settle in this local orbital. This process gives the appearance of the local spin having flipped from down to up, while conduction states have flipped from up to down, without any actual spin flipping. These two processes, the departure of the old local electron with spin down and the arrival of the new local electron with spin up, are totally independent and can occur in any order with any spin state. The average spin flipping would therefore appear to be random as up and down electrons come and go independently. This is also reflected in the formulation above that the spin index remains arbitrary and the energy is degenerated with respect to the spin of the electrons. This is correct however only if $U = 0$. Once $U$ becomes significant, the two processes are correlated. It would be energetically unfavourable to have both up and down spin electrons on the same local orbital, and so the second process would be more likely to occur after the first one. In this case, the local spin appears

Figure 2.2: The relative populations of conduction (c) and localized (f) electrons in the lower energy $E_-$ state $\Phi_-$. 

![Graph showing relative populations of c-electrons and f-electrons vs. $e_1$.]
Figure 2.3: The energy diagram of the non-perturbed system as a function of momentum $k$, with $t = 1, \epsilon_f = 1$ and $V = 2$. The gap between the two energy levels is exactly 4, the value of $U$.

to have flipped. The spin property of the model would therefore only show up in the perturbed Hamiltonian. We will come back to this case again in chapter 3 when we discuss the perturbation of a single site model.

In the following, the first and second order perturbation will be calculated based upon the ground state of the system as discussed. The correction in energy will be performed in the first and second order of $U$. The calculation is valid only if $U$ is small and hence the correction would not in general substantially alter the result of the ground state energy.

### 2.2.1 First and Second order perturbations

To find the first order perturbation, we simply apply the eigenstates of the unperturbed system to the perturbed term. Since the perturbation has two localized electron number operators, the result is non-zero only if the eigenstate contains two localized electrons in a single state. This can be accomplished by putting two quasiparticles of different spin orientations into one state without mutual inter-
2.2. Perturbative expansion of the small U limit

The energy diagram of the non-perturbed system as a function of $\varepsilon_f$, with $t = 1$, $k = 0$ and $V = 2$.

action, according to the solution of the unperturbed Hamiltonian. Since we are working in momentum space, the perturbed Hamiltonian would also need to be transformed accordingly.

$$\langle \Phi_{\pm \uparrow}(k) \Phi_{\pm \downarrow}(k) | n_{k\uparrow}^f n_{k\downarrow}^f | \Phi_{\pm \uparrow}(k) \Phi_{\pm \downarrow}(k) \rangle = \beta^4_{\pm}(k) \quad (2.29)$$

The first correction to the ground state energy of the system for a small but non-zero $U$ value is therefore,

$$\delta E^1_- = U \left( \frac{E_-(k) + 2t \cos(k)}{\sqrt{(E_-(k) + 2t \cos(k))^2 + V^2}} \right)^4 \quad (2.30)$$

The ground state in this case refers to the state with lowest overall energy, $\Phi_-(k)$.

In the second order perturbation, one would need to consider all the excited states of a ground state system. As $E_-$ is lower than $E_+$, $\Phi_{-\sigma}$ is the ground state and $\Phi_{+\sigma}$ the excited state. The second order perturbation process considers a quasiparticle from the ground state jumping up to the excited state for a short period of time and then returning to the ground state. Since more than one excited state is possible, the calculation will need to consider all possible excited states.
Fortunately there are only two.

\[
\langle \Phi_{+\uparrow} \Phi_{+\downarrow} | n_{k\uparrow}^f n_{k\downarrow}^f | \Phi_{-\uparrow} \Phi_{-\downarrow} \rangle = \beta_+^2 \beta_-^2 
\]
(2.31)
\[
\langle \Phi_{+\uparrow} \Phi_{-\downarrow} | n_{k\uparrow}^f n_{k\downarrow}^f | \Phi_{-\uparrow} \Phi_{-\downarrow} \rangle = \beta_+ \beta_-^3 
\]
(2.32)

The second order correction based upon these excited states is therefore

\[
\delta E_2 = U^2 \sum_{\psi \text{excited state}} \frac{(\langle \psi | n_{k\uparrow}^f n_{k\downarrow}^f | \Phi_{-\uparrow} \Phi_{-\downarrow} \rangle)^2}{E - E_\psi} 
\]
\[
= U^2 \left( \frac{\beta_+^2 \beta_-^2}{2E_--2E_+} + \frac{\beta_+^3 \beta_-}{E_--E_+} \right) 
\]
(2.33)

All basis states in this perturbation do not separate the up spin electrons from the down spin ones. As shown in the Eqs. (2.31) and (2.32), the basis for the perturbation must contain two localized electrons of different spin. Lacking any one of the two electrons would cause the perturbation to vanish. The system will therefore only be penalized, by raising the overall energy of the system for double occupation without lifting the degeneracy between up and down spin electrons. The magnetic property of the model in which \( U \) is small is therefore insignificant, compared with the resonance of the virtual bound state from the potential scattering. This explains why most of the models in the small Coulomb repulsion limit have insignificant spin interaction.

We turn to another approach in the next chapter by which the magnetic property of the model will be more pronounced from the calculation. In this approach, both \( U \) and \( V \) are assumed to be more substantial and the hopping matrix \( t \) is treated as a small parameter of the perturbation. This is the weakly hopping limit of the Anderson model, which is also called the single site limit.
CHAPTER 3

Perturbation Theory of the Single Site Limit

We have seen that the small $U$ limit does not provide the interaction we need from last chapter, another approach will be required and be the focus of this chapter. The second approach of approximation, as already mentioned at the beginning of last chapter, is to perturb the elements of the hopping matrix $t$. The unperturbed system in this case contains everything in the Hamiltonian except the kinetic energy term of the conduction electrons without which the system is equivalent to an array of single site model. This model, also named 'Anderson atom' [21], can be envisaged as a single atom in a solid or crystal structure in which the electron-electron interaction is identical to the one described in the Anderson model. We will, however, refer it as the single site Anderson model [65] to emphasis on the fact that the intersite interaction is ignored in the calculation.

The eigenstates of the unperturbed system calculated in this chapter use a method different to that of the last chapter, as the unperturbed Hamiltonian obviously cannot be diagonalized. The number of eigenstates is nevertheless limited and can possibly be found from the physics of the model.

The single site model in general can accommodate up to four electrons, including two opposite spin electrons in the conduction band (at Fermi energy level) and two opposite spin electrons at the impurity level. The model therefore has five different electronic states, with zero, one, two, three and four electrons at each site. The cases of zero and four electrons are of no interest to us, as no electron movement is allowed and the model is completely inert and non-interacting. Because of this, only the remaining three cases will be discussed below.

The Hamiltonian of the Anderson lattice can be rewritten as

$$H = T + \sum_i H_i$$  \hspace{1cm} (3.1)
3. Perturbation Theory of the Single Site Limit

where \( T \) denotes the kinetic energy of the system given by

\[
T = t \sum_{i,j,\sigma} c_{i\sigma}^\dagger c_{j\sigma}
\] (3.2)

\( H_i \) describes the rest of the terms in real space and the summation \( \sum' \) is restricted to the nearest neighbours only. The hopping matrix \( t \) is usually restricted to nearest neighbour anyway. The expression of \( H_i \) is reproduced here in real space, as opposite to momentum space used in Eq. (1.7).

\[
H_i = E_f \sum_{\sigma} f_{i\sigma}^\dagger f_{i\sigma} + U f_{i\uparrow}^\dagger f_{i\uparrow} f_{i\downarrow}^\dagger f_{i\downarrow} + V \sum_{\sigma} (c_{i\sigma}^\dagger f_{i\sigma} + f_{i\sigma}^\dagger c_{i\sigma})
\] (3.3)

\( E_f \) is the real space counterpart of \( \epsilon_f \) in momentum space. When the elements in the hopping matrix are small enough, one can use a Taylor expansion in The higher order terms of \( t \) will be treated as a perturbation to the solution of the stationary single-site model. First of all we will need to calculate the electron state of the system when \( t \) is zero.

### 3.1 Four electron states of a single atom

The calculation of the energy states and levels follows from the work of Bastide and Lacroix [12] in which the parameter \( U \) in the Hamiltonian was assumed infinite. This presumption, however, will not apply here because the most general solution of the energy level is being pursued. As mentioned before, a single site system can have five different electron states, of which only three states are of interest to us. We will discuss each one of them individually in the following section. Different to the last section, all calculations in this perturbation are performed in real space and no Fourier transform is required in the derivations.

#### 3.1.1 One-electron states

In one electron states the system contains only one electron, either a \( f \)-orbital localized electron or a conduction electron at site \( i \) of the lattice with spin \( \sigma \). The wave function of the state is therefore composed of two terms, one with \( c^\dagger \) and the other one with \( f^\dagger \). When \( f \)-electrons and \( c \)-electrons are completely isolated, they have their own energy levels determined by \( E_f \) and \( \epsilon_c \) in the Hamiltonian. However, when they are mixed together in a system, the state of the system changes and the \( \epsilon_c \) is in general non-zero, but if we take the Fermi energy level as our energy reference, \( \epsilon_c \) would be identically zero which is what we assume for the rest of the calculation.
3.1. Four electron states of a single atom

Energy levels vary. The lower energy state or the bonding state can in general be written as:

\[ |iS_{1,\sigma}^1\rangle = [\sin\theta c_{i\sigma}^\dagger - \cos\theta f_{i\sigma}^\dagger]|i0\rangle \]  

where \(|i0\rangle\) denotes the state of the lattice at site \(i\) with no electron occupation, and \(|iS_{1,\sigma}^1\rangle\) denotes the lowest energy state at site \(i\) with one electron of spin \(\sigma\).

Applying the Hamiltonian \(H_i\) to the eigenstate in Eq. (3.4) gives,

\[ H_i|iS_{1,\sigma}^1\rangle = [-V \cos\theta c_{i\sigma}^\dagger + (V \sin\theta - E_f \cos\theta) f_{i\sigma}^\dagger]|i0\rangle \]  

Equating this to the eigenstate in Eq. (3.4) with eigenvalue \(E_{S_{1}^1}\), i.e. \(E_{S_{1}^1}[(\sin\theta)c_{i\sigma}^{\dagger} - (\cos\theta)f_{i\sigma}^{\dagger}]|i0\rangle\), yields

\[ \frac{-V}{E_{S_{1}^1}} = \frac{E_f - E_{S_{1}^1}}{V} = \frac{\sin\theta}{\cos\theta} \]  

or

\[ E_{S_{1}^1}^2 - E_f E_{S_{1}^1} - V^2 = 0 \]  

The solution of this quadratic equation with the lowest energy eigenstate in Eq. (3.4) is

\[ E_{S_{1}^1} = \frac{E_f - \sqrt{E_f^2 + 4V^2}}{2} \]  

The coefficients of the \(c_{i\sigma}^\dagger\) and \(f_{i\sigma}^\dagger\) in the eigenstate of Eq. (3.4) can be found by substituting Eq. (3.8) back to Eq. (3.6),

\[ \sin\theta = \frac{1}{\sqrt{2}} \left(1 + \frac{E_f}{\sqrt{E_f^2 + 4V^2}}\right)^{\frac{1}{2}}, \]  

\[ \cos\theta = \frac{1}{\sqrt{2}} \left(1 - \frac{E_f}{\sqrt{E_f^2 + 4V^2}}\right)^{\frac{1}{2}} \]  

These two solutions are plotted in figure (3.1). The other solution of the quadratic Eq. (3.7) corresponds to the state with higher energy level. Substituting this solution back to Eq. (3.6) yields the new coefficients of \(c_{i\sigma}^\dagger\) and \(f_{i\sigma}^\dagger\) for the new eigenstate,

\[ |iS_{2}^1\sigma\rangle = [(\cos\theta)c_{i\sigma}^{\dagger} + (\sin\theta)f_{i\sigma}^{\dagger}]|i0\rangle \]  

and the corresponding energy of this state,

\[ E_{S_{2}^1} = \frac{E_f + \sqrt{E_f^2 + 4V^2}}{2} \]
The energy level of this state, together with the level of the lower energy state are plotted in figure 3.2. It shows the 'almost' crossing of two different energy lines, one with $E = 0$ and the other $E \propto E_f$. These two lines would have crossed if the conduction and the impurity electrons are not hybridizing. The hybridization between $f$-electrons and $c$-electrons allows for a mutual interaction, which results in a system with lower overall energy. This, as a consequence, creates a gap between the two energy levels at the point where they would have crossed, as shown in figure (3.2) and turn the model into a two band system. This gap has a dramatic consequence to the configuration of the system which will be described in the following. When $E_f$ is very small, the ground state of the system has energy proportional to $E_f$, indicating that the system is dominated by the localized $f$ electrons. In fact from the coefficients of the $f^\dagger$ and $c^\dagger$ terms in the wave function Eq. (3.4), one can see that the population of the localized $f$-electrons outweights the population of the conduction electrons in this lower energy state, as shown with dashed line in figure 3.1. On the contrary, when $E_f$ is very large, the energy level of the lower energy state approaches zero, as indicated on the right hand side of figure 3.2. The population
3.1. Four electron states of a single atom

![Graph showing two eigenenergies of a one-electron single site Anderson system. The solid line corresponds to the lower energy state while the dashed line to the first excited state. The parameters of the functions are the same as figure 3.1.](image)

Figure 3.2: The two eigenenergies of a one-electron single site Anderson system. The solid line corresponds to the lower energy state while the dashed line to the first excited state. The parameters of the functions are the same as figure 3.1.

of the conduction electrons, \( \sin^2 \theta \), in this state outweighs the population of the impurity electrons as opposed to the case when \( E_f \) is small, which indicates that most of the electrons in the system have turned to conducting states.

All of these results can also be understood from the fact that electrons are localized when \( E_f \) is very small and free when \( E_f \) is very large. This is because, when \( E_f \) is small, most of the conduction electrons hop down to the impurity site to gain the energy difference between the impurity and the conduction energy levels. Since the number of electrons can never exceed one, the double occupation interaction has no relevance to this state. Therefore, the ground state of the system can be approximated by the state with only impurity electrons. When \( E_f \) is very large on the contrary, conduction electrons require extra amount of energy to hop up to the impurity level and therefore remain at the conduction energy level as long as possible. As a result, the ground state of the system can be approximated by the state in which the conduction electrons dominate. These are confirmed by the relative density of the impurity and the conduction electron operators in the ground state at two extreme values of \( E_f \), as shown in figure 3.1.
3. Perturbation Theory of the Single Site Limit

When \( E_f \) is close to zero \((i.e., \) the conduction energy level\)), neither of the two extreme pictures of the ground state is correct. In this case, the conduction electrons and the impurity electrons coexist in the model and interchange through the hybridization process. It is through this process that the energy of the system is lowered and a gap is created between the two energy levels. Since no double occupation is possible in one electron state, this case is the same as the non-interacting case when \( U = 0 \) as described in section 2.2. This is sometimes referred as the ‘mixing valence’ regime when two states of the electrons coexist in the model. Interesting phenomena may arise when one state forms scattered islands in the sea of another state but this unfortunately is beyond the scope of this thesis.

### 3.1.2 Two-electron states

When two electrons are allowed in a single site Anderson Model, the electrons will arrange themselves into two different states, the singlet or the triplet, depending on the spin of the electrons. The singlet state of two electrons can be written as a combination of the product of two electrons, in such a way that the spin function of the state is anti-symmetric\(^1\). In terms of electron operators, it can be written as

\[
|iS^2\rangle = [\alpha c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger + \beta f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger + \gamma (f_{i\uparrow}^\dagger c_{i\downarrow}^\dagger - f_{i\downarrow}^\dagger c_{i\uparrow}^\dagger)]|0\rangle \tag{3.13}
\]

where \( \alpha, \beta \) and \( \gamma \) are normalized ratios of the three different electron configurations, and are different to the ones defined in Eq. (2.2).

Applying the interacting Hamiltonian in Eq. (3.3) to this state yields

\[
H_i|iS^2\rangle = [-2V\gamma c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger + (\beta(2E_f + U) - 2V\gamma)f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger + (-\alpha V - \beta V + \gamma E_f)(f_{i\uparrow}^\dagger c_{i\downarrow}^\dagger - f_{i\downarrow}^\dagger c_{i\uparrow}^\dagger)]|0\rangle \tag{3.14}
\]

Equating this result to \( E_{S^2}|iS^2\rangle \) gives three simultaneous equations in terms of the eigenvalue \( E_{S^2} \) of a two electron singlet,

\[
\alpha = -\frac{2V\gamma}{E_{S^2}}, \tag{3.15}
\]

\[
\beta = \frac{2V\gamma}{2E_f + U - E_{S^2}}, \tag{3.16}
\]

\[
\gamma = \frac{\alpha + \beta}{E_f - E_{S^2}}V \tag{3.17}
\]

\(^1\)In other words, the interchange of the spin of two electrons will recover the original state with a minus sign. Since the overall wave function, which is composed of the product of the spin wavefunction and the space wavefunction, for a set of two electrons must be anti-symmetric, the space function of the state must therefore be symmetric.
3.1. Four electron states of a single atom

There is one more equation to be satisfied by the \( \alpha, \beta, \gamma \), namely the orthonormality of the eigenstate.

\[
\alpha^2 + \beta^2 + \gamma^2 = 1 \tag{3.18}
\]

One can eliminate \( \gamma \) by substituting \( \alpha \) and \( \beta \) from Eqs. (3.15) and (3.16) into Eq. (3.18).

\[
E_{s_2}^3 - (3E_f + U)E_{s_2}^2 + ((2E_f + U)E_f - 4V^2)E_{s_2} + 2V^2(2E_f + U) = 0 \tag{3.19}
\]

Although this cubic equation can be solved by using the corresponding formula in [1], the solution is fairly complex. To get some insight into the solution without resorting to solving the cubic equation, we can look at two limiting cases of the equation and locate the range of energy within which the solutions will fall.

We begin with the variation of the parameter \( U \). When \( U = 0 \), Eq. (3.19) becomes

\[
(E_f - E_{s_2})(E_{s_2}(2E_f - E_{s_2}) + 4V^2) = 0 \tag{3.20}
\]

which requires either \( E_f = E_{s_2} \) or \( E_{s_2}^2 - 2E_fE_{s_2} - 4V^2 = 0 \). The solutions of these equations give the eigenvalue of \( H_i \) when \( U = 0 \),

\[
E_{s_2}^1 = E_f, \tag{3.21}
\]

\[
E_{s_2}^2 = E_f + \sqrt{E_f^2 + 4V^2}, \tag{3.22}
\]

\[
E_{s_2}^3 = E_f - \sqrt{E_f^2 + 4V^2} \tag{3.23}
\]

On the contrary when \( U \rightarrow \infty \), Eq. (3.19) becomes

\[
E_{s_2}^2 - E_fE_{s_2} - 2V^2 = 0 \tag{3.24}
\]

which yields only two solutions

\[
E_{s_2}^1 = \frac{E_f - \sqrt{E_f^2 + 8V^2}}{2}, \tag{3.25}
\]

\[
E_{s_2}^2 = \frac{E_f + \sqrt{E_f^2 + 8V^2}}{2} \tag{3.26}
\]

Based on these two limiting cases, as \( U \) increases from 0 to \( \infty \) the number of eigenvalues changes from three to two. This implies one of the three eigenvalues in the case of \( U = 0 \), goes to \( \pm \infty \) as \( U \) goes to infinity. By plotting the cubic equation with the value of \( U \) varying between 0 and infinity in figure 3.9, we found that it is the largest eigenvalue that diverges as \( U \) approaches \( \infty \). This agrees with the intuition that if \( U \) increases, the energy required for doubly occupying an impurity
site will also increase, discouraging two f-electrons from staying at the same site with opposite spin. In this case, the energy level of the excited state would therefore rise.

Another interesting observation from this cubic Eq. (3.19) is that it always bears three real and distinct solutions, given the positive definite values of the energy parameters $E_f$, $U$ and $V$. This follows from the procedure for solving the cubic equation using the formula given in [1].

After finding all the eigenvalues, the coefficients $\alpha, \beta, \gamma$ in the eigenstate can be evaluated for each one of these eigenvalues. It is easy to solve $\gamma$ from Eqs. (3.18), (3.15) and (3.16) to get

$$\gamma = (E_{S^2}(2E_f + U - E_{S^2}))(4V^2((2E_f + U - E_{S^2})^2$$

$$+ E_{S^2}^2 + 2E_{S^2}^2(2E_f + U - E_{S^2})^2)^{-1/2}$$

(3.27)

$$\alpha = -\frac{2V}{E_{S^2} \gamma}$$

(3.28)

$$\beta = \frac{2V}{2E_f + U - E_{S^2} \gamma}$$

(3.29)

To complete the solution, let's consider the other state, the triplet state, of the system with which by definition the spin function of the state must be symmetric whereas the space function is anti-symmetric. There are two states fulfilling this condition.

$$|i T^2_1\rangle = \frac{1}{\sqrt{2}}(f^\dagger_{i\uparrow}c^\dagger_{i\uparrow} + f^\dagger_{i\downarrow}c^\dagger_{i\downarrow})|0\rangle$$

(3.30)

$$|i T^2_{2,\sigma}\rangle = f^\dagger_{\sigma}c^\dagger_{\sigma}|0\rangle$$

(3.31)

Applying the interacting Hamiltonian in Eq. (3.3) to both of them gives

$$H_i|i T^2_1\rangle = \frac{1}{\sqrt{2}}E_f(f^\dagger_{i\uparrow}c^\dagger_{i\uparrow} + f^\dagger_{i\downarrow}c^\dagger_{i\downarrow})|0\rangle = E_f|i T^2_1\rangle,$$

(3.32)

$$H_i|i T^2_{2,\sigma}\rangle = E_f|i T^2_{2,\sigma}\rangle,$$

(3.33)

which yield only one eigenvalue $E_f$. Therefore, the triplet states have only one eigenenergy $E_f$ which is far above the lower singlet eigenenergy, but below the highest singlet antibonding energy. This is also shown in the energy level diagram in figure 3.9.

A similar analysis can be applied to the same model using the variation of $E_f$, instead of $U$, to obtain the dependence of the energy eigenvalue on $E_f$. The energy Eq. (3.19) in this case would have to be solved. We, nevertheless, chose to simulate the solution at different parameter regions and picked one that represents the best among all of the others. The explanation in this context is therefore based on
3.1. Four electron states of a single atom

Figure 3.3: The relative populations of $f^+f$ (dashed line), $c^+c$ (dash-dot line) and $f^+c + c^+f$ (solid line) electrons in a two electron single site Anderson Model vs $E_f$. The model is plotted with $U = 10$ and $V = 1$. This particular set of parameters, which nonetheless should be valid for any other parameter region. The parameters chosen in this case are $U = 10$ and $V = 1$. Taylor expansions were used from time to time to find the analytic expression of the solution around the corresponding fixed points.

We first begin with an extreme case, when $E_f$ is far below the conduction energy band, i.e. $E_f << -U$. In this case most of the conduction electrons would hop down to the impurity energy level with double occupation, and hence lower the total energy of the system. Since there are only two electrons in each site, the ground state of the model can be approximated by the doubly occupied impurity electron state which we denote $\phi_2$. This ground state puts the parameters $\alpha$ and $\gamma$ close to zero in Eqs. (3.28) and (3.28) but leaves $\beta$ in Eq. (3.29) close to one. The energy level of the ground state can hence be approximated by $2E_f + U$ for two electrons. After expanding Eq. (3.19) around this value of $E_{S2}$, we estimated the correction of the ground state energy from this first order term:

$$E_{S2} = 2E_f + U + 2V^2/(E_f + U) + O(V^4) \tag{3.34}$$
The first order correction turned out to be negative and so the actual ground state energy is lower than our asymptotic value. The impurity is nonmagnetic. This energy level is represented by the solid line on the right hand side of figure 3.4. The states of the electrons corresponding to this energy level are also shown on the right hand side of figure 3.3.

As $E_f$ increases and gets close to the value of $-U$, the model approaches the first mixed-valence region in which double occupation and single occupation of the impurity electron coexist in the model. When the hybridization term starts to diverge, the configuration of the ground state is beginning to change from double impurity occupation state $\phi_2$ to a single occupation state which we denote $\phi_1$. The second electron from the double occupation becomes energetically favourable to hop up to the conduction band and leaves only one electron behind in the impurity site. There is a region where the energy level $E_f + U$ is of the order of $V$ on either side the x-axis, (ie. $-U - V\sqrt{2} < E_f < -U + V\sqrt{2}$). In this region, the state with double impurity occupation and single conduction/impurity occupation is strongly mixed. When substituting $E_f = -V$ into Eq. (3.19), the ground state energy becomes...
3.1. Four electron states of a single atom

\[ E_{S^2} = E_f - V\sqrt{2} \] (3.35)

At this point, the model is approximately in the configuration \( \phi_{1,5} \) which has half double occupation and half single occupation of the impurity electrons. This can be verified from the calculation of parameters \( \alpha, \beta \) and \( \gamma \).

As \( E_f \) continues to increase, the density of impurity double occupation reduces and the ground state becomes mainly singly occupied \( \phi_1 \). When \( E_f \) is in the range of \(-U + \sqrt{2} < E_f < -V\sqrt{2}\), the ground state is dominated by \( \phi_1 \) and its energy is

\[ E = E_f + \frac{2V^2 U}{E_f(U + E_f)} \] (3.36)

The density of the electrons at different states is shown in the middle region of the solid curves in figure 3.3. The population of the double impurity occupation \( \phi_2 \) is decreasing as \( \beta \) approaches zero with increasing \( E_f \). On the contrary, the state \( \phi_1 \) with single occupation on both conduction and impurity site is growing, as \( \gamma \) increases to the peak at \( E_f = -U/2 \). The system at this point is in the symmetric case in which the lower and the upper impurity energy levels at \( E_f \) and \( E_f + U \) respectively, are sitting symmetrically about the conduction energy level. The energy gained from putting an electron in an empty impurity level is identical to the energy cost to doubly occupy an impurity site. As a consequence, the number of holes is equal to the number of electrons in the model. A hole in this case is defined as the state with three electrons.

After passing the symmetric case, the state with only single occupation \( \phi_1 \) is diminishing while the state with doubly occupied conduction electron \( \phi_0 \) is expanding. This is caused by the close proximity of the lower impurity level \( E_f \) to the conduction band, whereby the energy required to hop from the impurity level to the conduction band is substantially reduced.

Now \( \phi_1 \) is the ground state in this region of parameters. The ground state contains one f-electron and one c-electron. \( U \) is too large to permit the c-electron to fall into a \( f \) state, but \( E_f \) is negative enough to permit the f-electron to transfer to a conduction state. This is the region in which we would expect a magnetic moment to occur. Since the energy level of triplet state is simply \( E_f \), the energy difference between the singlet (ground state) and the triplet state, also called an effective exchange interaction, is

\[ E_{\text{singlet}} - E_{\text{triplet}} = \frac{2V^2 U}{E_f(U + E_f)} \] (3.37)

As \( E_f \) is negative and \( U + E_f \) is positive, this exchange interaction is negative (antiferromagnetic). This result was also found by Schrieffer and Wolff [78] who
performed a canonical transformation onto a single impurity Hamiltonian and obtained a spin flip term which corresponds to the magnetic moment of the model. The difference between the single impurity model and our single site model is merely the mobility of the conduction electrons which nevertheless is irrelevant as a static model is being considered. This explains the similarity between the results from these two models.

If \( E_f \) continues to increase to the conduction band at zero energy or within \( \pm V\sqrt{2} \), the ground state would be degenerated with the coexistence of the single particle \( f \) and \( c \) states and the double occupation of the \( c \)-electrons. This is the second mixed-valence region as opposed to the first mixed-valence region when \( E_f + U \) is closed to zero. When \( E_f \) is zero, the number of impurity electrons is approximately half, and the ground state energy is

\[
E = -\sqrt{2}V
\]

(3.38)

When \( E_f \) continues to rise above zero, the population of the impurity electrons reduces and the ground state is dominated by the conduction electrons. Most of the impurity \( f \) orbitals are now empty but not all of them because of the hybridization nature of the model. The number of impurity electrons is very small, approximately \( 2V^2/E_f^2 \). The ground state energy is

\[
E = -\frac{2V^2}{E_f}
\]

(3.39)

Both \( \beta \) and \( \gamma \) are close to zero now and \( \alpha \) is approaching one. This corresponds to the solid line on the right hand side of figure 3.4 when \( E_f \) is very large. The relative population of the electrons at different states is shown in figure 3.3. More details of this discussion can be found in [21].

After exploring the ground state behaviour, we can now turn to the first excited states. The first excited states are states which have their energy levels closest to the ground state energy level, represented by the dotted line in figure 3.4. These states can be found from solving the coefficients \( \alpha, \beta, \gamma \) in Eqs. (3.28,3.29,3.28) with \( E_f \) substituted by the second largest root of the third order Eq. (3.19). One example of these states is shown in figure (3.5) when \( U = 10 \) and \( V = 1 \).

When \( E_f < -U \), the ground state contains two localized electrons as previously discussed and the first excited state will contain predominantly an impurity electron and a conduction electron. This state is the closest state to the ground state in energy with the excitation energy being \( |E_F + U - V| \). The excitation energy is precisely the energy required for a localized electron to hybridize and hop to the conduction energy level. The populations of doubly occupied impurity electrons
3.1. Four electron states of a single atom

Figure 3.5: The relative populations of the three two-electron states in the first excitation vs $E_f$. The parameters used are the same as in figure 3.3

and conduction electrons are almost zero, as shown with the dashed line in figure 3.5.

As $E_f$ increases to the order of magnitude of $U$, the system is approaching the first mixed valence regime as discussed before. Double occupation of the localized site becomes less favourable in the ground state. In this case, it would cost more energy to doubly occupy an impurity site than singly occupying both the conduction and localized sites whenever $E_F > -U$, making the double occupying state the first excited state. Therefore, the first excited state starts to change from the single occupation state $\phi_1$ to the impurity double occupation state $\phi_2$.

The system at this point basically contains half of the electrons in $\phi_1$ and the other half in $\phi_2$.

When $E_f$ continues to rise and approach zero, the first excitation energy increases accordingly. The doubly occupied impurity electrons in the first excited state continues to increase to a peak just before $E_f$ reaches $-U/2$, as shown with the dotted line in figure 3.4. This is the state just before the symmetric case when $E_f = -U/2$. In the symmetric case, the conduction energy level is sitting in the middle of the impurity energy level $E_f$ and the impurity double occupation level
$E_f + U$. As discussed before, the ground state consists mostly of singly occupied conduction and impurity electrons. Since it costs the same amount of energy to doubly occupy an impurity site as to doubly occupy a conduction site, the first excited state contains either a double occupation of the conduction electrons or impurity electrons. Which one will dominate depends on the value of $E_f$ relative to the conduction energy level. For instance just before the symmetric case when the conduction energy level is closer to $E_f + U$ than $E_f$, a conduction electron is more energetically favourable to hop up to the doubly occupied impurity level $E_f + U$, than a localized electron to hop up to the conduction energy level. Therefore, double occupation of localized electrons will dominate in the first excited state. On the contrary when the conduction energy level is closer to $E_f$ than $E_f + U$, it will cost less energy for a localized electron to hop up to the conduction energy level, than a conduction electron to hop up to a doubly occupied impurity state. In this case, double occupation of conduction electron is more favourable energetically and will become the first excited state.

When $E_f$ continues to rise above the conduction energy level, the ground state contains mostly conduction electrons. Thus the first excited state will again contain
single occupation of conduction and impurity electrons with excitation energy $E_f + V$. This is the energy required to push a conduction electron from the ground state to the single impurity energy level. The entire process is shown in figure 3.5 in which the change of the population of electron states is clearly depicted with respect to the changes in $E_f$.

Beyond the first excitation, an electron can also be excited to the second excited state given enough excitation energy. The second excited states are states which are the second closest to the ground state in energy. Since the first excited state brings the system up to at least single occupation of either conduction or impurity site, the second excited states are most likely to contain double occupation of either conduction or impurity electrons as shown in figure 3.6. However, single occupation of conduction and localized sites are still possible in the symmetric case when the double occupation of localized electrons changes over to the double occupation of conduction electrons, as plotted in the middle of figure 3.6. The corresponding energy is the highest among the three excited states, as shown with the dashed line in figure 3.4.

### 3.1.3 Three-electron states

Similar to the one-electron states but with either f-electron or c-electron site fully occupied, the eigenstate with three electrons can be written as

$$|iS_1^3\sigma\rangle = [(\cos \xi)f_{i\sigma}^\dagger c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger - (\sin \xi)c_{i\sigma}^\dagger f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger]|i0\rangle$$  \hspace{1cm} (3.40)

It is interesting to note that this state is very similar to the one-electron state in Eq. (3.4). In fact, if we consider the four electron state as the ground state, this state corresponds to the one hole state.

Applying the interacting Hamiltonian in Eq. (3.3) to the eigenstate in Eq. (3.40) gives,

$$H_i|iS_1^3\sigma\rangle = [(E_f \sin \xi - V \cos \xi)f_{i\sigma}^\dagger c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger + (2E_f \cos \xi - V \sin \xi + U \cos \xi)c_{i\sigma}^\dagger f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger]|i0\rangle$$  \hspace{1cm} (3.41)

Equating it to the eigenstate in Eq. (3.40) multiplied by the eigenvalue $E_{S_1}$ of a three electron singlet,

$$E_{S_1}[(\sin \xi)f_{i\sigma}^\dagger c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger + (\cos \xi)c_{i\sigma}^\dagger f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger]|i0\rangle$$  \hspace{1cm} (3.42)

gives two simultaneous equations

$$E_{S_1} \sin \xi = E_f \sin \xi - V \cos \xi$$  \hspace{1cm} (3.43)
$$E_{S_1} \cos \xi = 2E_f \cos \xi - V \sin \xi + U \cos \xi$$  \hspace{1cm} (3.44)
These two equations can be combined by eliminating \( \tan \xi \), and become

\[
V^2 = (E_f - E_{s1})(2E_f + U - E_c)
\]

or

\[
0 = (E_f - E_{s1})^2 + (E_f - E_{s1})(E_f + U) - V^2
\]

The positive solution of this equation corresponds to the eigenvalue \( E_{s1} \) of the lower energy state, given by

\[
E_f - E_{s1} = \frac{-(E_f + U) - \sqrt{(E_f + U)^2 + 4V^2}}{2}
\]

and the corresponding coefficients of the electron operators in the eigenstate of Eq. (3.40) are

\[
\cos \xi = \frac{1}{\sqrt{2}} \left(1 + \frac{E_f + U}{\sqrt{(E_f + U)^2 + 4V^2}}\right)^{3/2}
\]

\[
\sin \xi = \frac{1}{\sqrt{2}} \left(1 - \frac{E_f + U}{\sqrt{(E_f + U)^2 + 4V^2}}\right)^{3/2}
\]

For the higher energy state, the energy level is given by the second solution of the quadratic Eq. (3.46). With this new energy level substituted back to Eqs. (3.43) and (3.44) for solving the coefficients of the electron operators in the eigenstate of Eq. (3.40), one gets

\[
| \text{i}S_2^3\sigma \rangle = [(\sin \xi) f_{i\sigma}^\dagger c_i^\dagger c_i^\dagger + (\cos \xi) c_i^\dagger f_{i\sigma}^\dagger] | \text{i}0 \rangle
\]

where \( \cos \xi \) and \( \sin \xi \) are given by Eqs. (3.48) and (3.49), and the energy eigenvalue \( E_{s2} \) of this state is

\[
E_f - E_{s2} = \frac{-(E_f + U) + \sqrt{(E_f + U)^2 + 4V^2}}{2}
\]

The relative population of f-electron and c-electrons shown in figure 3.7 and the energy levels in figure 3.8 are similar to the corresponding figures of the one electron states, except that the gap between the two levels as well as the cross-over point is now proportional to the parameter \( U \). The density of f-electrons and c-electrons are also very similar to those of the one electron states. The reason for the similarity is due to the particle-hole symmetry. In one electron state, an electron can be in any one of the four empty electron sites, whereas in three electron state a hole can be in any one of the four fully occupied electron sites. The energy of the states in the latter case however is different to the former, because of the on-site Coulomb
3.1. Four electron states of a single atom

interaction term $U$ which explains why the energy level in the latter case depends on $U$. Therefore, the same description as in the case of one electron will also apply here, only that a particle in a one electron state is substituted by a hole in three electron state.

Most of the results above are interpreted using the variation of $E_f$, with $U$ being held constant. It is, however, also possible to look at the model from another angle, with the help of figure 3.9. On the left of the figure when the model has only one electron and $E_f$ is assumed zero, the one-electron states split up into two eigenstates with eigenvalues $+V$ and $-V$ independent of the parameter $U$. This is because the Coulomb repulsion term in the Hamiltonian has no impact to a system with only one electron.

With two electrons and $U = 0$, as mentioned in the context there exists four eigenstates, three singlets of eigenvalues $-2V, 0, +2V$ and one triplet of eigenvalue 0. As $U$ increases, the three eigenvalues increase accordingly but within different extents. The lowest eigenvalue increases to the maximum $-\sqrt{2}V$ as $U \to \infty$ and the second eigenvalue increases to the maximum $+\sqrt{2}V$, while the highest one increases...
3. Perturbation Theory of the Single Site Limit

Figure 3.8: The two eigenenergies of a three-electron single site Anderson system. The parameter of the model is the same as figure 3.7.

to infinity. This increase in energy reflects the situation when two electrons hop to the same impurity site at the expense of extra energy. When the Coulomb repulsion $U$ is very large, the possibility of having two impurity electrons at the same site is almost zero, which forces the coefficient $\beta$ of $f^\dagger f^\dagger$ in the two-electron eigenstate of Eq. (3.13) to zero. This consequently transforms the eigenstate in Eq. (3.13) from having three coefficients to two and therefore reduces the energy levels from three to two.

When there are three electrons coexisting at the same site, the electrons behave in a similar way as to the one electron case. As shown in the figure when $U = 0$, the energy levels of the one electron case are identical to the energy levels of the three electrons one, since the on-site impurity interaction disappears when $U = 0$. With $U$ increasing from zero, the energy levels also increase. The negative energy level increases to the maximum zero, while the positive energy level soars up to infinity as $U$ goes to infinity. This rapid increase of energy level corresponds to the strong on-site interaction between two impurity electrons as $U$ becomes large. At the extreme case where $U$ goes to infinity, the large repulsive force forbids double occupation of impurity site and hence leaving only one eigenstate for each spin of
3.2. Perturbative expansion of the weak hopping limit

The first order perturbation of the static Anderson Model with respect to the hopping matrix elements $t$ provides the slow moving characteristics to the otherwise stationary system. The change in energy due to this slow electron motion varies with the density of electrons in the system. This can be attributed to the Coulomb interaction between the localized electrons which discourages two electrons of opposite spin hopping to the same site. It will be shown that the first order perturbation is not adequate to get the energy difference between the two spin configuration of

![Diagram](image_url)

**Figure 3.9:** The energy levels of one, two and three electrons in the single site impurity Anderson Model.

the electron. This agrees with the result that only one energy level survives as $U$ goes to infinity.

Both pictures, with the variation of $U$ and $E_f$, agree with each other. Both of them however only show the static behaviour of the model without any possible dynamical properties. To obtain the dynamic properties and magnetic characteristics of the model, one would need to go to the first and second order perturbation of the Hamiltonian with respect to the hopping matrix elements $t$ in the kinetic term $T$ of Eq. (3.2). The magnetic characteristics in particular can be derived from the energy difference between moving electrons of different spin orientations. This will be the focus of the next two sections.

3.2 Perturbative expansion of the weak hopping limit
electrons at the same site, the second perturbation is therefore needed.

Dimensionality will also have an effect on the perturbation since the possibility of hopping off from a particular site is higher with higher degrees of freedom. This essentially allows an electron to 'avoid' doubly occupying a localized orbital and hence reduces the ground state energy. To simplify the calculation, we would however only concentrate on a one dimensional lattice. The same calculation can, nevertheless, be extended to higher dimensions [12].

3.2.1 First Order perturbation

We will start from three different ranges of electron density \( n \).

For \( 0 < n < 1 \), less than one electron on average would be found in each site on average and the effect of the Coulomb interaction is negligible. Using the eigenstates given by Eq. (3.4) and extending them to two site model, the first order perturbation is

\[
\delta E^1_\sigma = \langle 0, S^1_{1,\sigma} | T | S^1_{1,\sigma}, 0 \rangle = t \sin^2 \theta
\]  

(3.52)

where \( \theta \) is defined in Eqs. (3.9) and (3.10) and the interaction between the two sites is negligible. This is the energy required to move a singlet ground state from one site to another when \( t \) is assumed to be small. It is in fact quite easy to interpret this result. As described in the last section, the nominal population of conduction electrons is \( \sin^2 \theta \). Among the conduction and the localized electrons, only the conduction electrons can hop from site to site, with the hopping matrix element \( t \). Therefore, the energy induced by hopping, among all electrons must be \( t \sin^2 \theta \).

When \( n = 1 \), there is one electron in each site on average. Any single hopping of an electron from this state would cause the state to accommodate more than one electron, no matter how short a life time that would be. In other words, it is impossible for the eigenstate of one electron system to jump to the same eigenstate of the system by a single hopping. This renders the first order perturbation of one electron per site zero.

\[
\langle S^1_{1,\sigma}, S^1_{1,\sigma'} | T | S^1_{1,\sigma'}, S^1_{1,\sigma} \rangle = 0
\]

(3.53)

This would also be true for systems of two, three and four electrons per site, as a single hopping will change the total number of electrons per site and therefore depart from their original states. It is, however, different in all other cases when the average number of electrons per site is not an integral number. The first order
3.2. Perturbative expansion of the weak hopping limit

The normalized second order perturbation energies (energy divided by $t^4$) vs $E_I$ impurity energy level when $U = 10$ and $V = 1$ and electron density $n$ equals 1/2 (solid line), 1 (dashed line) and 3/2 (dash-dot line).

Notice that the result is spin independent and so the orientation of the spin has no effect on the first order energy. The second order perturbation is needed to find the energy difference between states of different electron spins.

### 3.2.2 Second Order Perturbation

The second order perturbation is very different to the first order perturbation because of the extra intermediate state involved. Any state in the second order perturbation is allowed to go to an excited state before coming back to the same state. This on one hand increases the complexity of the calculation, but on the other hand opens up a great deal of possibility for a state to jump back to itself.
When $n = 1/2$, the number of excited states is limited to $|0, S_{1\sigma}^1\rangle$, the same as in the first order perturbation which is due to the limited number of electron in the state. The second order perturbation is

$$\delta E_{\sigma}^2 = (0, S_{1\sigma}^1)$$

in which we have defined

$$\langle A_{\sigma}, B_{\sigma} \rangle = \sum_{C,D\in \text{excited states}} \frac{(\langle C, D|T|A_{\sigma}, B_{\sigma}\rangle)^2}{E_A + E_B - E_C - E_D} \quad (3.59)$$

where $A_{\sigma}, B_{\sigma}, C_{\sigma}$ and $D_{\sigma}$ can be any state of the model. The excited state in the summation in this case involves only one state $\langle S_{1\sigma}^1 |$ for $n = 1/2$, as mentioned before and hence the perturbed energy is

$$\delta E_{\sigma}^2 = t^2 \frac{(\sin(\theta) \cos(\theta))^2}{E_f^2 + 4V^2} \quad (3.60)$$

When $n = 1$, the result is very different since the number of electrons is now twice as many as in the case of $n = 1/2$. The number of excited states has also increased dramatically which include: $|0, S_{2}^1\rangle, |0, S_{2}^2\rangle, |0, S_{2}^3\rangle, |0, T_{1}^2\rangle, |0, T_{2\sigma}\rangle, |S_{2}^1, 0\rangle, |S_{2}^2, 0\rangle, |S_{2}^3, 0\rangle, |T_{1}^2, 0\rangle$ and $|T_{2\sigma}, 0\rangle$.

The second order perturbation energy is therefore,

$$\frac{\langle S_{1\sigma}^1, S_{1\sigma}^1 |T|0, S_{1}^1\rangle^2}{2E_{S_{1}^1} - E_{S_{1}^2}} + \frac{\langle S_{1\sigma}^1, S_{1\sigma}^1 |T|0, S_{2}^1\rangle^2}{2E_{S_{1}^1} - E_{S_{2}^2}} + \ldots \quad (3.61)$$

This energy is plotted as a dashed line in figure 3.10 for the case of $U = 10$ and $V = 1$. The steep decrease of the curve near $E_f = 1$ is caused by one of the excited state energy levels $\langle 0, S_{1}^1 |$ in the summation Eq. (3.59) being too close to the ground state energy for two electrons. The difference in energy is $O(V^4)$ when $E_f \approx 0$ which causes the perturbation theory to break down. The result is reliable only up to the point near the early stage of the curve before it plummets down.

Having found the perturbation, we are now ready to find the difference in energy between excited states of electrons with different spin orientations. If the spins of the electrons at two neighbouring sites are anti-parallel,

$$\langle S_{1\sigma}^1, S_{1\sigma}^2 \rangle = 2\left( \sum_{x=1,2,3} t^2 (\alpha_x \sin^2 \theta + \beta_x \cos^2 \theta + \sin \theta \cos \theta \gamma_x)^2 \right) \frac{E_f - \sqrt{E_f^2 - 4V^2 - E_{S_{2}^2}}}{E_f - \sqrt{E_f^2 - 4V^2 - E_{S_{2}^2}}}$$

$$+ \frac{\sigma t (\sin(\theta) \cos(\theta)/\sqrt{2})^2}{-\sqrt{E_f^2 + 4V^2}} \quad (3.62)$$

$$+ \frac{\sigma t (\sin(\theta) \cos(\theta)/\sqrt{2})^2}{-\sqrt{E_f^2 + 4V^2}} \quad (3.63)$$
3.2. Perturbative expansion of the weak hopping limit

where \( \alpha_x, \beta_x, \gamma_x \) for \( x = 1, 2, 3 \) are the three solutions of Eq.3.28 for \( \alpha, \beta \) and \( \gamma \). However, if the spins of the electrons are parallel, there is only one excited state, namely \( T_\sigma \), that is permissible for the summation.

\[
(S_{1,\sigma}^1, S_{1,\sigma}^2) = 2 \frac{(t \sin(\theta) \cos(\theta))^2}{-\sqrt{E_f^2 + 4V^2}} 
\]  

(3.64)

This difference lifts the degenerated state energies between the parallel and anti-parallel electrons in the lattice and is exactly what we need to determine the magnetic properties of the model. One can also consider this energy difference as the result of an interaction between parallel and anti-parallel spin electrons. This interaction can be calculated by comparing the second order energy of two neighbouring particles and of two distant particles. The energy difference between two neighbouring particles with spins \( \sigma \) and \( \sigma' \) and the same but distant particles is given by

\[
(S_{1,\sigma}^1, S_{1,\sigma'}^1) - (S_{1,\sigma}^1, 0) - (S_{1,\sigma'}^1, 0) 
\]  

(3.65)

When \( \sigma = \sigma' \), this energy is identically zero as shown in Eqs.(3.60) and (3.64). This means that there is no interaction between particles of parallel spins when
3. Perturbation Theory of the Single Site Limit

\( n = 1 \). Figure 3.11 shows this energy with \( \sigma = \sigma' \) in dashed line and \( \sigma = -\sigma' \) in solid line. The sharp decrease of energy in the latter case is again due to one of the excited states having its energy too close to the energy of the ground state. One can see from this interaction that the lattice favours anti-ferromagnetism in the second order perturbation.

If one adds an extra electron per site to the model to make \( n = 3/2 \), the number of intermediate states in the second order perturbation increases even further and involves different kinds of excited states. Moreover, the Coulomb pair interaction starts to become significant as the number of double occupation increases. This double occupation is the result of having more than one electron in each site on average. The dash-dot line in figure 3.10 shows this second order energy. When \( E_f \) is far less than the conduction energy level at zero, doubly occupying the impurity sites is the best way to minimize the total energy and hence perturbation with respect to electron hopping yields an insignificant result. However, when the system is approaching the symmetric case at \( E_f = -U/2 \) (which is -5 in figure 3.10), it costs \( U/2 \) amount of energy for a conduction electron to doubly occupy a localized site. On the other hand a conduction electron would also gain the same amount of energy if it jumps to an empty localized site. In this case, conduction electrons have the freedom to remain unchanged or use the energy from occupying an empty localized site to doubly occupy another localized site. This resonating phenomenon of electrons allows conduction and localized electrons to interchange from one to another, and consequently lower the overall energy of the system as indicated by the minimum point in figure 3.10. Any imbalance of the energy level would break this symmetry and cause one kind of electron to dominate, which in turn increases the total energy.

At \( n = 2 \), the second order perturbation is exactly zero which happens to be identical to the first order perturbation. In this case, the electrons are basically static when the first and the second order perturbations with respect to electron hopping are zero. One can also find the spin interaction in the case of \( n = 2 \), by calculating

\[
(S^2_1, S^2_1) + (S^1_1, S^1_1) - (S^2_1, S^1_1) - (S^1_1, S^2_1)
\]

(3.66)

Since the first term is zero as mentioned before, this interaction can be calculated from the second and the third terms which we have already evaluated. The result is plotted in figure 3.12 when \( U = 10 \) and \( V = 1 \). It is in fact very similar to the result of \( n = 1 \) in figure 3.11, except for a hump at \( E_f = -U/2 \) caused by particle-hole symmetry in the symmetric case as explained before.

When \( n = 2.5 \) with the ground state being considered as \( |S^3_1, S^1_{1,\sigma}\rangle \), the ex-
Figure 3.12: The normalized second order perturbation energies (energy divided by $t^4$) vs $E_f$ impurity energy level with parallel spin $\sigma = \sigma'$ (dashed line) and anti-parallel spin $\sigma = -\sigma'$ (solid line) for electron density $n = 2$.

cited states include $|S_{(1,2),\sigma}^3, S_{(1,2),\sigma}^2\rangle$, $|S_{(1,2),\sigma}^3, T_{1}^2\rangle$, $|S_{(2,3),\sigma}^2, T_{2,\sigma}^2\rangle$, $|S_{(1,2),\sigma}^3, S_{(1,2),\sigma}^3\rangle$, $|T_{2,\sigma}^3, S_{(1,2),\sigma}^2\rangle$, $|T_{1}^2, S_{(1,2),\sigma}^3\rangle$. These states are very similar to the ones with $n = 1.5$ and the result of the perturbation calculation behaves similarly to that of $n = 1.5$. This is shown in the dash-dot line in figure 3.13 with $U = 10$ and $V = 1$, in which the minimum point has shifted towards $E_f = -U$. The magnitude of the curve is approximately twice as large as the $n = 1.5$ case, as the scale of the y-axis is twice that of figure 3.10. This increase in energy is mainly due to the Coulomb pair interaction among the localized electrons when the amount of pair occupation is greater with higher density of electrons. When $E_f$ is far below conduction level, every localized state is at least doubly occupied as there are more than two electrons per site on average. However when $E_f$ is increased to $-U$, the energy of a doubly occupied impurity electron is the same as a conduction electron and so conduction electrons can easily doubly occupy an impurity site and vice versa. This extra degree of freedom for electrons causes the overall energy of the system to drop to a minimum point at $E_f = -U$ as shown in figure 3.13.

When $n$ reaches three, there is only one vacancy at each site for electrons to
3. Perturbation Theory of the Single Site Limit

Figure 3.13: The normalized second order perturbation energies (energy divided by $t^4$) vs $E_f$ impurity energy level when $U = 10$, $V = 1$ and electron density $n$ equals $7/2$ (solid line), $3$ (dashed line) and $5/2$ (dash-dot line).

move to. This vacancy can also be considered as a hole in the sea of electrons, with higher energy content from the Coulomb interaction. The behaviour of the model is very similar to that when $n$ is equal to one. This can indeed be predicted from the similarities in the excited states of the state $|S_{1,\sigma}^3, S_{1,\sigma'}^3\rangle$, which contains $|S_{(1,2,3)}^2, S^4\rangle$, $|T_{2,\sigma}^2, S^4\rangle$, $|T_{1,\sigma}^2, S^4\rangle$, $|S^4, S_{(1,2,3)}^2\rangle$, $|S^4, T_{2,\sigma}^2\rangle$, $|S^4, T_{1,\sigma}^2\rangle$. $S^4$ represents the full electron state of a site. The similarity is also shown in the result of the perturbation drawn in dashed line in figure 3.13, compared with the one in figure 3.10. Again, the line plummets down at $E_f \approx -U$ when one of the excited states has energy very close to the energy of the ground state, which is the same as in the case of $n = 1$. Even the spin interaction between electrons next to each other is very similar to the $n = 1$ case. This interaction is formulated as

$$
(S_{1,\sigma}^3, S_{1,\sigma'}^3) + (S_{1,\sigma'}^3, S_{1,\sigma}^3) - 2(S_{1,\sigma}^3, S_{1,\sigma'}^3),
$$

(3.67)

equivalent to the energy difference due to the proximity of electrons with different spins. The result is plotted in figure 3.14, with the dashed line corresponding to $\sigma = \sigma'$ and the solid line to $\sigma = -\sigma'$. Both curves have a larger hump compared to the previous case in figure 3.12. It is mainly due to the first mixed-valence of the
3.3 Thermodynamic properties of the single site model

Since the eigenstates of the system have already been investigated with the corresponding eigenvalues, the calculation of the thermodynamics properties becomes...
straightforward. In this section, we will first introduce the canonical ensemble in which the single site system resides. Without any particle exchange to the thermo-bath, the system is isolated. Energies of the system and the thermo-bath can fluctuate individually, but the total energy must be constant. As the temperature of the system depends on the total energy, it becomes a constant in this case. There are hence two conserved quantities, one is the temperature and the other the total number of particles in the system.

In reality, this kind of isolation is however almost impossible to achieve. A more realistic model would be to consider the grand canonical ensemble in which the system is allowed to exchange energy and particles with the thermo-bath. In this way, the thermodynamic properties of the whole system can be derived based upon the calculation of a single site with different constraints. This however is only a statistical average which does not consider any coherent phenomena and strongly correlated effects. The result does nevertheless provide a basic understanding of a non-interacting system.

### 3.3.1 Properties of canonical ensembles

We first consider the single site system which is isolated from the environment with a constant number of particles. The energy of the system in this case can exchange with the thermo-bath, to attain a constant temperature. The total energy of the system and the thermo-bath must however stay constant. In the canonical ensemble, the probability that an eigenstate remains in the system is inversely related to the eigenenergy of the state. In other words, the larger the eigenenergy of the state, the less probable the state will be found in thermodynamic equilibrium. The weighting factor according to Boltzmann’s distribution is

\[ p_i \propto \exp(-\beta E_i) \quad (3.68) \]

for any eigenenergy \( E_i \) in the state indexed \( i \) with \( \beta = 1/k_BT \). As the distribution is normalized, the proportionality factor can be written as

\[ Z = \sum_i \exp(-\beta E_i) \quad (3.69) \]

which is known as the canonical partition function.

The thermodynamic average of any variable is therefore given by

\[ \langle X \rangle = \frac{\sum_i X \exp(-\beta E_i)}{\sum_i \exp(-\beta E_i)} \quad (3.70) \]
3.3. Thermodynamic properties of the single site model

| state $\Phi$ | coeff. of $\Phi$ relative to $|0\rangle$ | energy $E_{S_i}$ | $\langle \Phi | S_i^z | \Phi \rangle$ | $\langle \Phi | S_i^z S_i^{tot} | \Phi \rangle$ |
|--------------|------------------------------------------|------------------|-----------------|-----------------|
| $S_{1,1}^t$  | $c_t \sin \theta - f^i_{\sigma} \cos \theta$ | $E_{S_{1,1}}$ | $\sigma_1^2$ | $\frac{1}{4} \cos^2 \theta$ |
| $S_{1,2}^t$  | $c_t \sin \theta - f^i_{\sigma} \cos \theta$ | $E_{S_{1,2}}$ | $\sigma_1$ | $\frac{1}{4} \sin^2 \theta$ |
| $S_{2,1}$    | $\alpha_1 c_1^1 c_1^1 + \beta_1 f^1_{\sigma} f^1_{\sigma}$ | $E_{S_{2,1}}$ | 0 | 0 |
| $S_{2,2}$    | $\alpha_2 c_1^1 c_1^1 + \beta_2 f^1_{\sigma} f^1_{\sigma}$ | $E_{S_{2,2}}$ | 0 | 0 |
| $S_{2,3}$    | $\alpha_3 c_1^1 c_1^1 + \beta_3 f^1_{\sigma} f^1_{\sigma}$ | $E_{S_{2,3}}$ | 0 | 0 |
| $T_{2,1}$    | $\frac{1}{\sqrt{2}} (f^1_{\sigma} c_1^1 + f^1_{\sigma} c_1^1)$ | $E_F$ | 0 | 0 |
| $T_{2,2}$    | $f^1_{\sigma} c_2^1$ | $E_F$ | $\sigma_1$ | $\frac{1}{4}$ |
| $S_{2,1}^t$  | $f^t_{\sigma} c_2^1 \cos \xi - c_2^1 f^t_{\sigma} \sin \xi$ | $E_{S_{2,1}}$ | $\sigma_1^2$ | $\frac{1}{4} \cos^2 \xi$ |
| $S_{2,2}^t$  | $f^t_{\sigma} c_2^1 \sin \xi + c_2^1 f^t_{\sigma} \cos \xi$ | $E_{S_{2,2}}$ | $\sigma_1^2$ | $\frac{1}{4} \sin^2 \xi$ |

Table 3.1: Summary of the eigenstates and eigenenergies of one, two and three electron states

where $X$ is any thermodynamic variable. In terms of the Hamiltonian $\mathcal{H}$, we can turn the summation of different energy levels $i$ into different eigenstates $\nu$ of the Hamiltonian.

$$\langle A \rangle = \frac{\sum_\nu \langle \nu | A \exp(-\beta \mathcal{H})| \nu \rangle}{\sum_\nu \langle \nu | \exp(-\beta \mathcal{H})| \nu \rangle}$$ \hfill (3.71)

For instance, the average number of localized electrons in the one, two and three electron system can be calculated from the formula

$$\langle n_f \rangle = \frac{\sum_\nu \langle \nu | n_f | \nu \rangle \exp(-\beta E_\nu)}{\sum_\nu \exp(-\beta E_\nu)}$$ \hfill (3.72)

in which $E_\nu$ corresponds to the eigenvalue of the Hamiltonian at eigenstate $\nu$, given the same number of total electrons.

Using table 3.1 the summary of the results from last section, we found for one electron,

$$\langle n_f \rangle = \frac{\cos^2 \theta \exp(-\beta E_{S_1}) + \sin^2 \theta \exp(-\beta E_{S_2})}{\exp(-\beta E_{S_1}) + \exp(-\beta E_{S_2})}$$ \hfill (3.73)

two electrons

$$\langle n_f \rangle = \frac{\sum_{i=1,2,3} (2 \beta_i^2 + \gamma_i^2) \exp(-\beta E_{S_{2,1}}) + 3 \exp(-\beta E_F)}{\exp(-\beta E_{S_1}) + \exp(-\beta E_{S_2}) + \exp(-\beta E_{S_3}) + 3 \exp(-\beta E_F)}$$ \hfill (3.74)

and three electrons

$$\langle n_f \rangle = \frac{(\cos^2 \xi + 2 \sin^2 \xi) \exp(-\beta E_{S_{3,1}}) + (\sin^2 \xi + 2 \cos^2 \xi) \exp(-\beta E_{S_{3,2}})}{\exp(-\beta E_{S_{3,1}}) + \exp(-\beta E_{S_{3,2}})}$$ \hfill (3.75)

The three different localized electron densities were plotted in figure 3.15 which shows typical thermodynamic behaviours. At high temperatures, all states of the
Figure 3.15: The densities of localized electrons in our model with (i) one-electron per site (solid), (ii) two-electron per site (dashed) and (iii) three-electron per site (dotted). They are plotted using $E_F = -2.5, V = 1$ and $U = 8$.

system are equally probable, and therefore the density of the localized electrons is half the density of the total number of electrons in the system. In the one electron case, it is 0.5; in the two electron case, it is 1; and in the three electron case, it is 1.5. The figure clearly shows the asymptotic behaviour of the three different cases to their limiting values at high temperature. When the temperature decreases, the system settles in the state with minimum amount of energy. In the three electron case, this would be the state with only one localized electron, to avoid the strong Coulomb repulsion. In the one and two electron cases, the lower energy state would contain a localized electron due to the negative value of $E_F$ to help reduce the total energy of the system. The behaviours are shown in the same figure with temperature close to zero.

Our next task is to find the magnetic susceptibility of this system. As it stands, it is impossible to do so since the Hamiltonian is independent of any external magnetic field. We therefore suppose that an external magnetic field $B$ is applied to the system and it couples only to the spin of the electrons. This brings an additional
3.3. Thermodynamic properties of the single site model

The Zeeman term, to the Hamiltonian,

$$H_B = g\mu_B\mathbf{B} \cdot \sum_i s_i,$$  \hspace{1cm} (3.76)

where $s_i$ is the spin operator of the $i$th electron, $g$ the Landé factor and $\mu_B$ Bohr’s magneton, the basic unit of magnetic moments of electrons. Without loss of generality, we assume the direction of the magnetic field is either parallel or anti-parallel to the spin of the electrons.

The total energy of the system is therefore,

$$E(B) = E(0) + g\mu_B m_B,$$  \hspace{1cm} (3.77)

where $m = -1/2, 1/2$ for one and three electron systems but $-1, 0, 1$ for two electrons. $E(0)$ is the energy of the eigenstate of the system when the magnetic field is absent. The magnetic field in the formulation is always assumed to be small, so the impact on the eigenstate of the original Hamiltonian is negligible. Indeed a small magnetic field is all we need to calculate the magnetic susceptibility.

Incorporating this new Zeeman term into our Hamiltonian, we can now calculate the magnetic moment, or magnetization, induced by the magnetic field $\mathbf{B}$.

$$M = g\mu_B \sum_i \frac{e^{-\beta E_i(B)} \langle i | S_z | i \rangle}{\sum_i e^{-\beta E_i(B)}}$$  \hspace{1cm} (3.78)

in which the index $i$ represents different eigenstates of the system with the same number of electrons. Using the results from last section which is summarized in table 3.1, the magnetic moment for one, two and three electrons are shown as follows:

$$M_1 = \frac{g\mu_B}{2} \tanh \frac{1}{2} g\mu_B B \beta,$$  \hspace{1cm} (3.79)

$$M_2 = \frac{2g\mu_B e^{-\beta E_F} \sin g\mu_B B \beta}{e^{-\beta E_{2,1}} + e^{-\beta E_{2,2}} + e^{-\beta E_{2,3}} + e^{-\beta E_F} + 2e^{-\beta E_F} \cosh g\mu_B B \beta},$$  \hspace{1cm} (3.80)

$$M_3 = \frac{g\mu_B}{2} \tanh \frac{1}{2} g\mu_B B \beta,$$  \hspace{1cm} (3.81)

The magnetic susceptibility by definition is the first partial derivative of magnetization with respect to $B$ when $B \to 0$.

$$\chi = \lim_{B \to 0} \frac{\partial M}{\partial B},$$  \hspace{1cm} (3.82)

This provides a direct way of obtaining the magnetic susceptibility from the results above. The magnetic susceptibilities of the system with one, two and three electrons
are therefore given by:

\begin{align}
\chi_1 &= \frac{1}{4}(g\mu_B)^2\beta \\
\chi_2 &= \frac{2(g\mu_B)^2\beta}{e^{-\beta(E_{s_{2,1}}-E_F)} + e^{-\beta(E_{s_{2,1}}-E_F)} + e^{-\beta(E_{s_{2,1}}-E_F)}} + 3 \\
\chi_3 &= \frac{1}{4}(g\mu_B)^2\beta 
\end{align}

(3.83) (3.84) (3.85)

The inverse of these susceptibilities are plotted in figure 3.16 as a function of temperature $k_B T$. The one electron and the three electron susceptibilities happen to be identical and are shown overlapping each other in the figure. The reason for this is simply because the systems in both cases are paramagnetic and the corresponding eigenenergies becomes degenerated with respect to up-spin and down-spin states. No difference will therefore be made with different orientation of electrons. This is indeed the famous Curie law for paramagnetic substances, which says the magnetic moment is always proportional to the applied field. The Curie's constants are given by the inverse of the slope in figure 3.16 which assumes different values for different materials. Paramagnetism was explained by Langevin [59] as existing only in those atoms which possessed a permanent magnetic moment. The applied magnetic field succeeded in aligning them against thermal fluctuations. The magnetic moment in this case originates from the spin of electrons.

In the two electron case, the susceptibility is quite different. At high temperatures, each state of the system is equally probable and so the energy is degenerated with different spins of electrons, similar to the one and three electron cases. As the temperature decreases, the electrons would try to singly occupy all the $f$-electron sites, to reduce the overall energy. This is the local moment regime. Since the energy gain from hybridizing to the local orbitals is spin independent for single occupations, the system still behaves in a paramagnetic fashion. As the temperature falls further, the system begins to discriminate among the states, based on their energy content. The local moment begins to disappear and the system starts to settle into its lowest singlet state. The susceptibility drops rapidly to zero as the temperature continues to reduce. This is due to the fact that the magnetic moment reduces to zero smoothly in the singlet state, independent of the applied magnetic field and therefore the derivative in the susceptibility is zero.

Although all our calculations were based on one site only, the results are still correct for many sites since the particles between sites are not interacting with each other directly. The electron from one site has no influence on an electron from any other site, according to our stationary Hamiltonian. The system is therefore an ideal ensemble. Each individual site of a many site lattice forms a system by itself
and at a given time occupies a certain single-site microstate. All other particles from other sites can be considered as providing the heat bath for each site in the lattice of particles at a given temperature.

It is important that electrons between different sites of a lattice are non-interacting. In other words, the probability of finding a particle in a certain microstate is completely independent of the microstates of particles at any other site. In ensemble theory, the individual single system of an ensemble must be independent of each other. With an interacting system this is no longer true as the microstate of a certain particle depends on those of other particles at some point in time and the result will derivates from what we obtained above.

Our calculation would have been enough if we were only interested in an isolated system of a fixed number of particles. When the actual number of electrons is unknown but only the observable average is provided, we have to turn to the grand canonical ensemble.

Figure 3.16: The inverses of magnetic susceptibility of electrons in our model with (i) one-electron per site (solid), (ii) two-electron per site (dashed) and (iii) three-electron per site (dotted) which coincides with (i). They are plotted using $E_F = -2.5, V = 1$ and $U = 8$. 
3.3.2 Properties of grand canonical ensembles

The difference between Grand Canonical Ensemble and Canonical Ensemble is that in Grand Canonical Ensemble, heat or energy as well as particles can exchange with the surroundings. In this kind of open system, only the observable average number of particle is fixed, the actual number of particles at any given time is in general different. This is very similar to the behaviour of energy in Canonical Ensembles in which energy could fluctuate across different parts of the system as well as over time, but the observable average of temperature must stay constant.

This concept of Ensemble is closer to reality. The actual number of particles is unknown to us; only the average is observable. Also, the exact energy in any part of the solid at any given point in time is not directly measurable. Only the average energy is accessible.

The implication of open systems to our model is that any site in the lattice may instantaneously contain any number of electrons from zero to four, depending on the chemical potential and the average number of electrons. Also, the susceptibility is more likely to be due to localized electrons, instead of all electrons as conduction electrons are free to exchange between different sites. The total spin of them will average out.

To consider our model in the Grand Canonical Ensemble, we replace the Hamiltonian by the grand Hamiltonian

$$\mathcal{H} = H - \mu N,$$

where $N$ is the number operator for the entire system, and $\mu$ is the chemical potential. $\mu$ has to be determined by the requirement that the thermodynamic average of $N$, $\langle N \rangle$, has a specified value. The thermodynamic average in Eq. (3.71) will now use $\mathcal{H}$ instead, which means that contributions from states with all possible numbers of particles $n$ are included, and are further weighted by a factor $e^{\beta \mu}$.

$$\langle X \rangle = \frac{1}{Z_G} \sum_i \langle i | X e^{-\beta (H + g\mu B \cdot B - \mu N)} | i \rangle$$

where $i$ is any one of the states of the system and $Z_G$ is the grand partition function

$$Z_G = \sum_i \langle i | e^{-\beta (H + g\mu B \cdot B - \mu N)} | i \rangle$$

We start the calculation with the average number of particles $\langle N \rangle$, to determine the chemical potential $\mu$.

$$\langle N \rangle = 2 + \frac{1}{Z_G} \left( \sum_{j=1,2} e^{-\beta (E_{S1,j} - \mu)} + \sum_{j=1,2} e^{-\beta (E_{S2,j} - 3\mu)} \right)$$
3.3. Thermodynamic properties of the single site model

Let \( z = e^{\beta \mu} \) called the fugacity, we can simplify this equation to a quadratic form of \( z \) multiplied by another \( z \).

\[
(2\langle N \rangle - 1) \sum_{j=1,2} e^{-\beta E_{S_{1,j}}} + (\langle N \rangle - 2)(\sum_{j=1,2,3} e^{-\beta E_{S_{2,j}}} + 3e^{-\beta E_F})z^2 \\
+ 2(\langle N \rangle - 3) \sum_{j=1,2} e^{-\beta E_{S_{5,j}}}z^2 = 0
\]

Non-trivial solutions exist only when

\[
(\sum_{j=i,1,2,3} e^{-\beta E_{S_{1,j}}} + 3e^{-\beta E_F})^2(\langle N \rangle - 2)^2 - 16(\sum_{j=1,2} e^{-\beta E_{S_{1,j}}})
(\sum_{j=1,2} e^{-\beta E_{S_{5,j}}})(\langle N \rangle - 1)(\langle N \rangle - 3) \geq 0
\]

or in general if \( \langle N \rangle \geq 1 \) and \( \langle N \rangle \leq 3 \), as this will guarantee the positivity of the LHS of the inequality. When \( \langle N \rangle < 1 \) or \( \langle N \rangle > 3 \), the inequality would never be satisfied and hence \( z = 0 \) is the only solution. This implies that the model becomes highly unstable and tries to drop the chemical potential in order to attract more electrons or holes to fill up the vacancy. This is only possible with the grand canonical ensembles as particles can exchange freely between the ensemble and the thermo-bath. This case nonetheless is out of the scope of our study.

Figure 3.17: The densities of localized electrons in our model using grand canonical ensemble with three different \( \langle N \rangle \). i) \( \langle N \rangle = 1.1 \) (solid), ii) \( \langle N \rangle = 2.1 \) (dashed), iii) \( \langle N \rangle = 2.9 \) (dotted). They are plotted using \( E_F = -2.5, V = 1 \) and \( U = 8 \).
3. Perturbation Theory of the Single Site Limit

Figure 3.18: The magnetic susceptibilities of electrons in our model using grand canonical ensemble with three different \( \langle N \rangle \). i) \( \langle N \rangle = 1.1 \) (solid), ii) \( \langle N \rangle = 2.1 \) (dashed), iii) \( \langle N \rangle = 2.9 \) (dotted). They are plotted using \( E_F = -2.5, V = 1 \) and \( U = 8 \).

The solution of the quadratic equation provides the value of the chemical potential, as a function of \( \langle N \rangle \) and \( \beta \) when \( \langle N \rangle \geq 1 \) and \( \langle N \rangle \leq 3 \). In the symmetric case when \( \langle N \rangle = 2 \) and \( U = E_F/2, \mu \) is always zero which simplifies the calculation significantly. This is however not our objective to obtain a result only for a single particular case, and hence we will continue to assume an asymmetric case in general.

Using the solution of \( \mu \) given the average number of particles \( \langle N \rangle \) and \( \beta \), we can calculate the average number of localized electrons from,

\[
\langle n_f \rangle = \frac{1}{Z_G} \sum_i \langle i | n_f e^{-\beta (H - \mu N)} | i \rangle
\]  

(3.92)

The plot of this equation, based on the summarized result in table 3.1, is shown in figure 3.17. This plot is quite similar to the corresponding plot with canonical ensemble. At high temperature when all states are equally probable, the density of the localized electrons is half the total density of electrons. As the temperature reduces, the system tries to reduce the total energy by going to states with single localized electron occupation. The rest of the electrons will therefore be left in the conduction band.
3.4 Remarks

We now calculate the susceptibility of the localized electrons in the presence of a magnetic field, starting from the magnetic moment $M$.

$$M = g\mu_B \langle S_z^f \rangle = \frac{1}{2} \langle (n_i^f - n_i^c) \rangle$$

$$= \frac{1}{2Z_G} \sum_i \langle i | (n_i^f - n_i^c) e^{-\beta(H+g\mu BS+\mu N)} | i \rangle$$

(3.93)

The first derivative of the magnetic moment $M$ with respect to the magnetic field when $B$ is zero gives the magnetic susceptibility

$$\chi_f = \beta(g\mu_B)^2 \langle S_z^f S_z^{tot} \rangle,$$

(3.94)

where $S_z^{tot}$ is the $z$ component of the total electron spin, $S_z^{tot} = \frac{1}{2} (N_\uparrow - N_\downarrow)$.

Plots of susceptibility for different values of $\langle N \rangle$ are shown in figure 3.18 which turns out to be quite different to that of the canonical ensemble. At high temperatures, the system is at the free-atom limit, similar to that of the canonical ensemble. When the temperature reduces, the system behaves in a paramagnetic fashion whereby most of the localized sites are singly occupied with two fold spin degeneracy. What is interesting is that this behaviour continues on even when the temperature drops to zero. Looking into the equation, we indeed found that at low temperature most of the electrons try to settle into a paramagnetic three electron state with two c-electrons and one f-electron. This behaviour is vastly different to the case when three electron states are projected out of the model in the approach of Bastide and Lacroix [12]. In the latter case, the behaviour of the magnetic susceptibility is very similar to that of the two-electron state in the canonical ensemble as only single impurity occupation is allowed in every site of the lattice. The system is therefore well into the local moment region and the magnetic susceptibility shows a ferromagnetic characteristic. This is the result of the system trying to achieve maximum amount of occupation.

3.4 Remarks

Perturbation is an excellent tool when the nature of the ground state is known. As shown in the last section, it can reveal the small energy difference between the parallel and the anti-parallel electron states which are degenerated in the first order calculation. Higher order perturbation can also be done in a similar fashion to find out more details of the energy difference. If we however started with a Hamiltonian of which the unperturbed part does not contain the essential ingredient to lift the spin degeneracy, perturbation will not produce the right results as clearly shown
in the perturbation in section 2.2. The reason is that the ground state of the unperturbed Hamiltonian is completely spin-independent, and the perturbed part is symmetrical between up and down spins. No energy difference would therefore be found between up and down spin electrons no matter how many orders of perturbation calculations were done. It is very similar to the case when perturbation is used to find out the self-energy of a potential scattering model. Without summing the perturbation series to infinite order, the pole of the Green's function corresponding to the self-energy would never move.

To sum the perturbation series to infinite order in our case is very difficult, if not impossible. Every other electron in the model provides an extra excited state for the higher order perturbation. Projective operators in the canonical perturbation theory however does provide an alternative to calculate the perturbation series to infinite order. This will be the focus of next chapter.
CHAPTER 4

Perturbation Theory of the Atomic Limit

The perturbative approach in the last chapter provides some interesting results, but its application is limited since highly correlated systems normally have a strong interactive nature between particles, and individual particles do move around either coherently or randomly. Alternative perturbation approaches with small hybridization $V$, should provide a more realistic picture of materials. The Periodic Anderson Model (PAM) can be solved exactly when $V$ is zero. No spin interaction is possible in this non-perturbed case, as no preference is made to either the up or the down spin. When $V$ is non-zero and $U$ is large, the picture is entirely different. In the tight-binding formulation of the Anderson lattice, the interaction between the particles is strong. As pointed out in Anderson's paper [6], the key parameters of this model are the Coulomb interaction $U$ and the charge-transfer energy $\Delta = \epsilon_f - \epsilon(k)$ which is the energy cost to transfer an electron from the conduction band to a localized site. In the study of high temperature superconducting oxide (HTSO), the latter would be the energy cost to transfer a hole from the Cu site to the O site.

Zaanen and Oleš [97] concluded, based on various spectroscopic data of the HTSO, that most of these compounds are in the regime of $U > \Delta$. Even then, there are still two possibilities for the configuration of the HTSO. The first one is when $V^2$ is small, giving rise to a correlation gap between the subbands which moves the compound to the class of charge-transfer semiconductors. The second case is when $\Delta$ is small enough that most of the lower two bands are filled and the system is well into mixed-valence itinerant regime. Both of these cases are interesting and will be discussed in the context. Zaanen and Oleš expanded the Hamiltonian of the periodic model with small $V$ based upon the canonical perturbation theory developed by Chao, Spałek and Oleš [23] initially for the Hubbard model. This provided them with the relative strength of electron interactions at different orders of $V$ and
the possibility of a pairing mechanism in these high temperature superconducting oxides.

Another approach from a totally different angle is given by Haldane [41] who expanded the partition function of the single impurity Anderson Model at the small $V$ atomic limit, using 'imaginary time' interval instead of energy. This method can be easily generalized to the tight binding periodic Anderson Model. So far, we are not aware of anyone having taken this path yet. Both of these approaches are in fact complimentary and hence will be discussed in detail in this chapter.

Cluster formation within a CuO$_2$ plane in the atomic limit using a perturbative approach was a controversial subject. The cluster is formed by four Oxygen holes surrounding a central Cu$^{2+}$ ion with phase consideration in the electronic orbitals. Zhang and Rice [98] using hybridization as in the PAM have shown that the binding energy of this cluster moving together in the bipartite lattice is close to a singlet bonding state the cluster. In order words, the cluster, commonly known as the Zhang-Rice singlet, seems to be energetically stable according to the perturbation calculation. We will discuss in more detail the formation of this state and its weakness in the next section.

In the following, we will first discuss the Zhang-Rice singlet using the Rayleigh-Schrödinger perturbation theory introduced in chapter 2. It is followed by the approach of Chao, Spalek and Oleš [23], hereafter denoted CSO, to introduce the canonical perturbation expansion theory (CPT), and then re-derive the result of Zaanen and Oleš [97] using diagrams. The diagrammatic expansion of the perturbative series using CPT is a new result of this thesis which provides a much more
intuitive way of understanding the underlying physics. Finally we will calculate the thermodynamic properties of the lattice model using Haldane's partition function expansion.

4.1 The Zhang-Rice Singlet

The formation of a pair or a singlet cluster in the high temperature superconducting cuprate has always been controversial. This singlet, if found stable, could lead to the explanation of high temperature superconductivity, similar to the 'Cooper pair' in the normal theory of superconductivity. To show that a pair or cluster is stable in its formation, the ground state energy (or the binding energy) of the cluster must be very close to the ground state energy of the lattice. Zhang and Rice [98] calculated the cluster of four Oxygen holes surrounding a Copper hole in the process of searching for an one-band effective Hamiltonian. They found that the binding energy of this cluster is very close to the singlet ground state of the lattice. The key for the formation of the singlet from their explanation is the hybridization which strongly binds a hole on each square of Oxygen ions to the central Cu$^{2+}$ ion. This singlet can move through the lattice of the Cu$^{2+}$ ions in a similar way as a hole does in a single band effective Hamiltonian. The fact that two singlet clusters residing on the same square strongly repelling each other suggests that the effective Hamiltonian must resemble a single band Hubbard model.

In order to keep the formulation consistent with the rest of this chapter, we will consider electrons, instead of holes, in the formation of the singlet. The calculations and the principle are essentially the same, except the charge of the carriers in consideration being opposite. The Hamiltonian in Eq. (1.7) must be written differently in this case to that of Zhang and Rice [98], but the physics remains the same. For convenience, the Hamiltonian is reproduced here again in real space and is separated into two parts $H = H_0 + H_V$ in which

$$H_0 = \sum_{i,\sigma} (\epsilon_c c_{i,\sigma}^\dagger c_{i+1,\sigma} + \text{H.C.}) + \sum_{i,\sigma} \epsilon_f f_{i,\sigma}^\dagger f_{i,\sigma} + U \sum_i f_{i,\uparrow}^\dagger f_{i,\downarrow}^\dagger f_{i,\downarrow} f_{i,\uparrow}$$

(4.1)

The definitions of the operators remains the same as in Eq. (1.7), except that they are now in real space. The energy of the $c$-electrons is a constant $\epsilon_c$, instead of $k$ dependent as in $\epsilon(k)$. The perturbed Hamiltonian is given by:

$$H_V = \sum_{i,\sigma} \sum_{l \in \{i\}} V_{il} f_{i,\sigma}^\dagger c_{l,\sigma} + \text{H.C.}$$

(4.2)

The summation on $l$ runs over the four nearest neighbour sites at the four corners of a square, centering at site $i$. The hybridization term has considered the phases
of the electronic wave functions and is written as

\[ V_{IL} = (-1)^{M_{IL}} V. \]  

(4.3)

\( V \) is the amplitude of the hybridization and is assumed to be very small compared with \( U, \epsilon_c, \epsilon_f \) and \( U - \epsilon_c \). \( M_{IL} \) is either 1 or 2, depending on the relative location of the indices \( i \) and \( l \).

\[ M_{IL} = \begin{cases} 2 & \text{if } l = i - \frac{1}{2} \vec{x} \quad \text{or} \quad i - \frac{1}{2} \vec{y} \\ 1 & \text{if } l = i + \frac{1}{2} \vec{x} \quad \text{or} \quad i + \frac{1}{2} \vec{y} \end{cases}. \]  

(4.4)

In other words, the phase factor is positive if the central f-electron is hybridizing with the left or the bottom c-electron and the factor is negative if the f-electron is hybridizing with the top or the right c-electron, as shown in figure 4.2.

The first step is to choose a proper set of c-electrons which would lead to an antiferromagnetic ground state, the ground state we found from the result of the weak hopping limit in chapter 3. The combinations of the four c-electron states together with the central f-electron form either a symmetric or anti-symmetric state with respect to the central f-electron. This is believed to provide the desired magnetic ground state for the model.

\[ P_{i,\sigma}^{(S,A)} = \frac{1}{2} \sum_{i \in \{i\}} (\pm 1)^{M_{i,l}} c_{i,\sigma}, \]  

(4.5)

where \((- (+)\) corresponds to the symmetric (S) and anti-symmetric (A) space function of the state\(^1\). The phase of c-electrons and f-electrons are defined in figure 4.2 in which a variable \( g \) is used to differentiate the location of the c-electrons. The factor \((-1)^g\) also provides the appropriate phase factor for the electrons residing in any one of the four corners. The symmetric and anti-symmetric states may combine with f-electrons to form either singlet or triplet states.

\[ \phi_{\text{singlet}}^{(S,A)} = \frac{1}{\sqrt{2}} (P_{i,\uparrow}^{(S,A)} f_{i,\downarrow} - P_{i,\downarrow}^{(S,A)} f_{i,\uparrow}); \]  

(4.6)

\[ \phi_{\text{triplet}}^{(S,A)} = \frac{1}{\sqrt{2}} (P_{i,\uparrow}^{(S,A)} f_{i,\downarrow} + P_{i,\downarrow}^{(S,A)} f_{i,\uparrow}). \]  

(4.7)

The energy content of the unperturbed Hamiltonian \( H_0 \) with respect to these two different states is simply

\[ \langle \phi_{i,\downarrow}^{(S,A)} | H_0 | \phi_{i,\uparrow}^{(S,A)} \rangle = \epsilon_d + \epsilon_p, \]  

(4.8)

\(^1\)As mentioned in a footnote of chapter 3, the total electronic wave function must be anti-symmetric. Since the wave function of a system of electrons is the product of the space and the spin wave functions, either one of the space or the spin wave functions must be anti-symmetric, but not both. The symmetric and anti-symmetric state here is referred to as the space wave function.
4.1. The Zhang-Rice Singlet

The Zhang-Rice Singlet

Figure 4.2: The electronic configuration of $CuO_2$ in oxide cuprates, in which the four Oxygen reside at four corners of the square while the $Cu^{2+}$ ion is at the centre. The $+$ and $-$ represent the phase factor of the wave functions and the $g$ variable identifies the location of the particles at any one of the four corners. Note: $(-1)^g$ gives the phase product at any one side of the $Cu^{2+}$ ion.

after substitution. It merely states that there is one f-electron and one c-electron in both the singlet ($s$) and the triplet ($t$) states. The first order perturbation of both the singlet and the triplet states using the Rayleigh-Schrödinger perturbation theory introduced in chapter 2 yield identical zero, because a single hybridization process applied to these states does not cause them to return to the original states. In other words,

$$\langle \phi^{(S,A)}_{s,t} | H_\nu | \phi^{(S,A)}_{s,t} \rangle = 0, \quad (4.9)$$

The energy splitting between the singlet and the triplet states must therefore be in the higher order perturbation. To calculate the second order perturbation as explained in section 2.1 of chapter 2, we need to know all the excited states of the unperturbed Hamiltonian. One of the excited states associated with the singlet and the triplet states are states with two c-electrons at any one of the four corners of
Theor y of the Atomic Limit

Figure 4.3: All possible excited states of a singlet cluster in the second order perturbation theory. The four diagrams correspond to four different possible excited states of the single symmetric ground state of a square cluster. The arrows on the circle are c-electrons resonating between four different corner sites of a square cluster. The solid arrows show one possible configuration, whereas the grey dashed arrows show other possibilities.

the square cluster as shown in diagram (a) of figure 4.3. These two c-electrons can resonate together between the four possible locations to lower the excitation energy. This excited state is labelled \( \psi_1^{(S,A)} \) depending on whether the state is symmetric or anti-symmetric.

The second excited state \( \psi_2^{S_{1,2}} \) is found when the two c-electrons after hybridization become nearest neighbour and resonate together, as shown in diagram (b) of the figure 4.3. The behaviour of this excited state is very similar to that of the previously mentioned excited state, except that this c-electron pair spans two sites instead of one single site as in the first case. This excited state is always symmetric with two different electron order sequences, differentiated by a superscript of \( S \).

Likewise, the third excited state \( \psi_3^{A_{1,2}} \) now has two c-electrons sitting at the two opposite diagonal ends of the central f-electrons after hybridization. The two electrons resonate together in this collinear fashion, as shown in diagram (c) of figure 4.3. This state, as opposed to the previous state, is always anti-symmetric.

The last excited state \( \psi_4 \) has two f-electrons doubly occupying the same site \( i \), as shown in diagram (d) of figure 4.3. This last excited state has no distinction between the symmetric and the anti-symmetric state as it occupies only one single site. The superscripts of \( (S,A) \) are hence dropped.

We first start with the symmetric ground state of the singlet and triplet configuration and consider all the possible excited states based on them. With the symmetric state, the first excited state can be written as

\[
\psi_1^A = \frac{1}{2} \sum_{g=1,2,3,4} c_g \psi g_{i}^1,
\]

in which the phase factor in the original ground state in Eq. (4.5) is compensated by the phase factor of the hybridization term in Eq. (4.3) and the \( g \)-factor is defined in
4.1. The Zhang-Rice Singlet

Figure 4.2. The symmetric ground state is therefore turned into an anti-symmetric excited state. The energy associated with the perturbation of this excited state is

$$\frac{\langle \psi_{\text{single}}^S | H_V | \psi_1^A \rangle^2}{E_{\psi_{\text{single}}} - E_{\psi_1}} = -\frac{2V^2}{\epsilon_c - \epsilon_f}, \quad (4.11)$$

The second excited state according to diagram (b) of figure 4.3 has the form of

$$\psi_2^S = \frac{1}{2} \sum_{g=1,2,3,4} (-1)^{g+g+2} c_{g,\uparrow} c_{g+2,\downarrow}, \quad (4.12)$$

for two neighbouring c-electron resonance. The sign factor is determined based on the assumption that the second c-electron in the equation is created from the hybridization of the central f-electron. The result would remain the same if we had chosen the first electron instead of the second one to undergo the hybridization, due to the indistinguishability of the two c-electrons. The behaviour of this excited state is very similar to that of the first state as mentioned previously, if one considers the two c-electrons as a pair. Unlike the first state, the phase factor in this case is not positive definite but dependent on the position of the electrons. The energy associated with this excited state is therefore

$$\frac{\langle \psi_{\text{single}}^S | H_V | \psi_2^S \rangle^2}{E_{\psi_{\text{single}}} - E_{\psi_2}} = -\frac{2V^2}{\epsilon_c - \epsilon_f}, \quad (4.13)$$

The third excited state is formed when the two electrons are at the two diagonal ends of the square cluster. This state can be written in the form

$$\psi_3^A = -\frac{1}{\sqrt{2}} \sum_{g=1,3} c_{g,\uparrow} c_{g+1,\downarrow}, \quad (4.14)$$

The negative sign is due to the opposite phase factors at any two diagonal ends of the square cluster. The excited energy associated with this state is

$$\frac{\langle \psi_{\text{single}}^S | H_V | \psi_3^A \rangle^2}{E_{\psi_{\text{single}}} - E_{\psi_3}} = -\frac{4V^2}{\epsilon_c - \epsilon_f}, \quad (4.15)$$

The last excited state of the symmetric singlet is the doubly occupied f-electron state $\psi_4 = f_{i,\uparrow} f_{i,\downarrow}$, which has the excitation energy of

$$\frac{\langle \psi_{\text{single}}^S | H_V | \psi_4 \rangle^2}{E_{\psi_{\text{single}}} - E_{\psi_4}} = -\frac{8V^2}{U - \epsilon_c - \epsilon_f}, \quad (4.16)$$

The total binding energy of the symmetric singlet state is therefore the sum of all the possible excitation energy above,

$$-8V^2 \left( \frac{1}{\epsilon_c - \epsilon_f} + \frac{1}{U - \epsilon_c - \epsilon_f} \right), \quad (4.17)$$
relative to the ground state energy. The excited states of the symmetric triplet ground state are essentially the same as those of the symmetric singlet ground state. The only difference is the opposite sign factor in the second term of Eq. (4.6) when the binding energy is calculated. Due to the anti-symmetrical nature of the spin functions in the space symmetric state, this sign factor cancels out exactly two equal terms in the equation. The excitation energy is therefore identically zero for the symmetric triplet state.

The next step of our consideration is the singlet and the triplet anti-symmetric ground states. The calculation of the excitation energy is very similar to that of the symmetric states, except for different excited states. Since the phase factor in the square cluster is removed from the anti-symmetric state, the sole cause of sign flip in this case must be from the hybridization. In other words, the sign compensation that we mentioned above for the first excited state would not happen here and therefore the first excited state in the anti-symmetric case is different.

\[
\psi_1^A = \frac{1}{2} \sum_{g=1,2,3,4} (-1)^g c_{g,\uparrow} c_{g,\downarrow},
\]

This change of sign however does not change the excitation energy.

Likewise, the excitation energy of the second excited state remains unchanged. The excited state itself however has a slightly different configuration to its counterpart in the symmetric case.

\[
\psi_2^{S_2} = \frac{1}{2} \sum_{g=1,2,3,4} (-1)^g c_{g,\uparrow} c_{g+2,\downarrow},
\]

Without loss of generality, this sign factor is again determined from the assumption that the second c-electron is hybridized from the central f-electron.

The difference between the symmetric and the anti-symmetric states arise from the third and the fourth excited states. In the third excited state, the hybridization puts the two c-electrons at the two diagonal ends of the square cluster. Depending on which electron of a resonating state is hybridized from the central f-electron, two identical results of opposite sign can be found from two different configurations of the same resonating state. This sign difference causes a cancellation of two equal contributions and hence reduces the excitation to zero. The same picture can also be applied to the fourth excited state when a doubly occupied impurity state can be excited by hybridizing a c-electron from two different configurations of the same resonating state, except with a sign difference. The total excitation energy of the anti-symmetric singlet is therefore

\[
-4V^2 \frac{1}{\epsilon_c - \epsilon_f}.
\]
4.1. The Zhang-Rice Singlet

The excitation energy of the anti-symmetric triplet state is identically zero for the same reason as for the symmetric triplet state.

As pointed out by Zhang and Rice [98], the large binding energy of the symmetric singlet state is due to the phase coherence. This phase coherence is achieved by assigning the same phase factor that appears in the square cluster to the hybridization of the Hamiltonian. In fact the binding energy of the whole cluster should also be compared with the binding energy of each CuO bond. Since the resonating effect is absent in a single bond, the binding energy to the second order perturbation is merely

\[ -2V^2 \left( \frac{1}{\epsilon_c - \epsilon_f} + \frac{1}{U - \epsilon_c - \epsilon_f} \right) \]  

(4.21)

which is the result of two doubly occupied excited states, one in the Oxygen site and the other Cu\(^{2+}\) site. The factor of two is from the spin degeneracy of the hybridization. This energy turned out to be only a quarter of the binding energy of the cluster. It is therefore extremely energetically favourable to form a square cluster based upon this result. This result is also consistent with our initial assumption that the ground state is anti-ferromagnetic, as the cluster state \( \psi^{S}_{\text{singlet}} \) is anti-ferromagnetic.

The calculation so far has only reached the conclusion that the ground state is a singlet and anti-ferromagnetic. We still have not shown that the cluster would move as a group together in the singlet state, and also the clusters would repel each other if they overlap. This repulsion is the equivalence of the Coulomb interaction between electrons or between holes in the Anderson model and the Hubbard model. We will first show the second point and then come back to the first one.

When a square cluster of c-electrons overlays another cluster of c-electrons with a central f-electron \( P_{i+}^{S} P_{j+}^{S} J_{i,j} \), the excited states of the combined clusters are very similar to the ones of the single cluster we discussed above, with the difference of an extra c-electron. The four excited states of this ground state is drawn in the diagrams (a) to (d) of figure 4.4. The diagrams (a) and (b) show a double occupied c-electron state in the nearest neighbourhood of a singly occupied c-electron. The extra electron is the result of hybridization from the central f-electron. These two excited states are in fact mirror images of each other and hence are not identical. The third state in diagram (c) shows a doubly occupied c-electron sitting at the opposite diagonal end of a singly occupied c-electron. Different to the corresponding diagram (c) in figure 4.3, the two sites of the c-electrons in this case are distinguishable and so the resonating state contains four different configurations, instead of just two in the former case. The last excited state in diagram (d) has a doubly occupied impurity site and one c-electron resonating at four corners around the impurity.
One interesting difference between these excited states and the previous ones in the single cluster case is that all four excited states of an overlapping cluster can resonate with four different configurations. The four excited states can be put in the form:

\[
\begin{align*}
\varphi^{(1)}_{i,\sigma} &= \sum_{g=1,2,3,4} c_{g,\uparrow} c_{g,\downarrow} c_{g+2,\sigma}, \\
\varphi^{(2)}_{i,\sigma} &= \sum_{g=1,2,3,4} c_{g,\sigma} c_{g+2,\uparrow} c_{g+2,\downarrow}, \\
\varphi^{(3)}_{i,\sigma} &= \sum_{g=1,2,3,4} c_{g,\uparrow} c_{g,\downarrow} c_{g+1,\sigma}, \\
\varphi^{(4)}_{i,\sigma} &= \sum_{g=1,2,3,4} f_{i,\uparrow} f_{i,\downarrow} c_{g,\sigma}.
\end{align*}
\]  

The corresponding excitation energies from the second order perturbation

\[
\frac{\sum_{\sigma} \langle P_{i,\uparrow}^S P_{i,\downarrow}^S d_{i,\sigma} \rangle H_V \varphi^{(1,2,3,4)}_{i,\sigma}}{E_{P_{i,\uparrow}^S P_{i,\downarrow}^S d_{i,\sigma}} - E_{\varphi^{(1,2,3,4)}}}
\]  

are shown here in the same order \(-2V, -2V, -2V, -4V\). The total binding energy of two overlapping clusters is therefore \(-V^2(6/(\epsilon_c - \epsilon_f) + 4/(U - \epsilon_c - \epsilon_f))\) which is much higher than the energy of two separated cluster states. As a result, the two clusters of electrons feel a strong repulsion when moving towards the same site. This shows that the cluster would move similarly to an electron or a hole in a single band model.

The picture so far obtained from the calculations is static. In fact the cluster might be stable when stationary, but as soon as the electrons acquire enough energy to move they may disperse and move in different directions. The final step is therefore to find if the individual electrons of a cluster gain energy when they move
4.1. The Zhang-Rice Singlet

together. Since moving together can be considered as having similar momentum amongst the group of electrons in the cluster, one could project, from many different combinations of movement of a cluster in momentum space, the movement that contributes to coherence displacement. To this end, we can simply Fourier transform the cluster of electrons to the momentum space and apply a small window function to it, so that only those electrons which are travelling with a similar momentum would remain in the wavefunction.

\[ P_{k,\sigma} = N_s^{-1/2} \beta_k \sum_i P_{i,\sigma}^S e^{-i\mathbf{k} \cdot \mathbf{R}_i}, \]  

(4.27)

in which \( N_s \) is the number of square clusters in the lattice and \( \beta_k \) is the window function which can also be considered as a factor that normalizes the commutation of the cluster operators \( P_{i,\sigma}^{(S,A)} \) due to the common corners shared between neighbouring clusters. Since we only want to consider electrons travelling with a very small momentum in one direction, this window function must be directional, periodic with a cycle of \( 2\pi \) in momentum space, large for small momentum \( k \) but small for \( k \) close to \( \pi \). One of the candidates chosen by Zhang and Rice [98] is the inverse of the raised cosine function in two dimensional space,

\[ \beta_k = \frac{1}{\sqrt{1 - (\cos k_x + \cos k_y)/2}}. \]  

(4.28)

Other possibilities are the periodic Gaussian function and the periodic Lorentzian function after proper normalization. These two functions would however introduce a first order discontinuity at the point where \( k \) is \( \pi \), as Gaussian and Lorentzian functions are non-periodic in their own nature.

The new singlet cluster of the \( c \)-electrons after an inverse Fourier transform becomes orthogonal and is complete in the \( c \)-electron space, when the triplet state is assumed negligible because of a large energy difference. This new cluster function in real space

\[ p_{i,\sigma} = N_s^{-1/2} \sum_k P_{k,\sigma} e^{i\mathbf{k} \cdot \mathbf{R}_i} \]  

(4.29)

can form a singlet or triplet with a \( f \)-electron at site \( i \) of a cluster.

\[ \chi_{i}^{s,t} = \frac{1}{\sqrt{2}}(p_{i,\uparrow}f_{i,\downarrow} \pm p_{i,\downarrow}f_{i,\uparrow}) \]  

(4.30)

where the positive sign is for triplet \( (t) \) and negative sign for singlet \( (s) \).

The excitation energy of this new square singlet cluster moving cooperatively in one direction was shown by Zhang and Rice [98] very close to the binding energy of the stationary counterpart. It is \( -8(1 + \lambda^2)t \) for singlet and \( -8(1 - \lambda^2)t \) for
triplet, in which $\lambda = N_s^{-1} \sum_k \beta_k^{-1} \approx 0.96$. The calculation was done based on the assumption that $\epsilon_c - \epsilon_f = U - \epsilon_c - \epsilon_f$ and using a new symbol $t = V^2/\epsilon_c - \epsilon_f$. Because of the large energy difference between the symmetric singlet and any other states described above including the triplet, the system can be treated completely within the singlet $\{\chi^s_i\}$ subspace.

Using these results, one can describe the square cluster of c-electrons as a quasiparticle which forms a singlet with the f-electrons at the centre of the cluster. Two quasi-particles repel each other if they reside at the same site, but otherwise travel freely within the lattice. This is just the single band Hubbard model in terms of quasi-particles in the large $U$ limit. The original paper [98] contains a more detail discussion of the relationship between the PAM and the Hubbard model which we have only gone through briefly. The paper also showed that the effective Hamiltonian can be related to the $t - J$ model.

The possibility of existence for this kind of singlet is indisputable, as the observation of the Zhang-Rice singlet has been claimed by many experiments. However, it would be an entirely different story if this Zhang-Rice singlet was the true ground state of the high temperature superconducting cuprate. To date, there are still many doubts about the explanation of the high temperature superconductivity using the Zhang-Rice singlet. In particular, the calculations described above ignore any interactions between clusters. The next nearest neighbour hopping is neglected even though it is believed to be significant in many superconducting cuprates. Only one particular cluster configuration has been singled out in the theory, but in reality there are many other configurations which could yield a similar binding energy. The competition between these different configurations could be significant. The real-space renormalization-group approach has already cast significant doubt on the stability of the singlet state [97], in relation to other new phenomena such as strips and phase separation.

In the next section, we look at the atomic limit of the Periodic Anderson Model from a completely different angle, using canonical perturbation theory. This approach provides a more realistic picture to the hopping of the carrier electrons in the lattice and some ideas about the spin interaction between electrons that are far apart. The phase factor introduced in this section is crucial when one considers the resonance bond theory, but since the focus for the rest of the thesis is rather on canonical perturbation theory, this phase factor will be ignored for simplicity.
4.2 Canonical Perturbation Theory

The canonical perturbation expansion that Chao, Spalek and Oleš (CSO) developed for the Hubbard Model basically divides the Hamiltonian into energy levels using projectors, and transforms the Hamiltonian canonically to remove the first order term of the inter-subband coupling. Only the lowest order terms of the canonical expansion are evaluated and this provides a good estimate of the ground state energy. This approach is very similar in some ways to the approach of the Schrieffer and Wolff [78], with perturbation technique included. CSO applied this theory to the Hubbard Model whereas here we will consider it for the Period Anderson Model. Instead of quoting the results, we will derive them from first principles.

In order to use the theory in its own form, we first need to project the Hamiltonian of the Anderson lattice into different energy subbands. Let us first define the projection operators to do so. With the Hubbard Model,

$$P_l = \sum_{i_1} \sum_{i_2 > i_1} \sum_{i_3 > i_2} \cdots \sum_{i_l > i_{l-1}} \left( \prod_{i \neq i_1 \ldots i_l} (1 - n_{i\uparrow} n_{i\downarrow}) n_{i_1 \uparrow} n_{i_2 \uparrow} \cdots n_{i_l \uparrow} \right)$$

(4.31)

is the projection operation of an sites lattice which projects out from a many-electron state, of which each particular configuration contains \( l \) doubly occupied sites. \( n_{i\sigma} \) is the number operator associated with the localized state at site \( i \) and spin \( \sigma \). In the Hubbard Model, each double occupation of a localized site raises the energy level of the system by a Coulomb interaction \( U \), similar to the corresponding term in the Anderson model. These operators divide the energy spectrum of the Hamiltonian evenly with each subband separated by a distance of \( U \), as shown in Figure 4.5.
For the Anderson lattice, the energy band structure is not as simple as that of
the Hubbard model, which is shown in figure 4.6. As mentioned above, the key
parameter of the Anderson model is $\Delta$ and $U$ which define the separation of each
subband in the energy spectrum above the conduction band. The upper subband is
$U - \Delta$ from the Fermi level, whereas the lower subband is $\Delta$ below it. We assume
in figure 4.6 that $U > \Delta$ for clarity but in general it does not need to be the case.
Every transition must change the energy by either $\pm(U - \Delta)$ or $\pm\Delta$. Unlike the
Hubbard model, the distance of the upper subband $(U - \Delta)$ from the conduction
band is generally different to that of the lower subband $\Delta$. A combination of multiple
upper and lower subbands will therefore cause multiple fine splitting of the upper
and the lower subbands, with respect to the Fermi level. This fine splitting is due
to the difference of $U - \Delta$ and $\Delta$. When $U - \Delta = \Delta$, the fine splitting disappears
and the energy spectrum becomes the same as that of the Hubbard model. In the
simple case when there is only one upper and one lower band for the excitation of
a single localized site, we can easily define the projection operators.

$$ P_0 = \sum_{i'} \left( \sum_{i \neq i', \sigma} \hat{n}^f_{i, \sigma} \right) (1 - n^f_{i', \sigma})(1 - n^f_{i', \downarrow}), \quad (4.32) $$

$$ P_1 = \prod_{i, \sigma} \hat{n}^f_{i, \sigma}, \quad (4.33) $$

$$ P_2 = \sum_{i'} \left( \sum_{i \neq i', \sigma} \hat{n}^f_{i, \sigma} \right) n^f_{i', \uparrow} n^f_{i', \downarrow}, \quad (4.34) $$
in which $P_0$ projects one empty localized site out of the whole lattice. $P_1$ projects
with single occupancy on all the localized sites and $P_2$ projects on the subband with
double occupancy of a localized site. $\hat{n}^f_{i, \sigma}$ is defined as the short hand notation of
$n^f_{i, \sigma}(1 - n^f_{i, -\sigma})$. In the Hubbard Model, electrons can hop between these subbands
freely, starting from the ground state which is normally the lowest energy in the
diagram. In the Anderson lattice model, the hopping is slightly more complicated.
As it is usually assumed that the ground state contains purely singly occupied
localized sites, electrons can hop to the next higher subband $P_2$ at the expense of
$U - \Delta$ in energy. On the other hand, localized electrons can hop up to the conduction
band to create an empty localized site which is $\Delta$ in energy above the ground state.
Either of these can be the first excited state of the model, depending on the relative
magnitude of the energy involved. Another way to look at the latter hopping is to
describe it in terms of holes, instead of electrons. In this case, a hole hops down to
the next lower subband $P_0$ and creates an empty localized site. The results are the
same, but this interpretation allows us to always start from the conduction band
and finish at the conduction band. To clarify this distinction, all hoppings in the
energy diagram that we show in the following will carry a letter (either h or e) to indicate which particle is involved. As the energy involved in holes hopping is always the negative of energy involved in electrons hopping, a minus sign \(-(...)\) is always shown with the energy in the diagram, as in figure 4.7 for example.

For the excitation of two localized sites, we would need to introduce more projection operators to cater for different energy subbands.

\[
P_{00} = \sum_{i \neq i'} \left( \prod_{\sigma, \tau, \tau'} \hat{n}_{i, \sigma}^f \right) (1 - n_{i', \tau}^f)(1 - n_{i', \tau'}^f)(1 - n_{i, \tau}^f), \quad (4.35)
\]

\[
P_{02} = \sum_{i \neq i'} \left( \prod_{\sigma, \tau, \tau'} \hat{n}_{i, \sigma}^f \right) (1 - n_{i', \tau}^f)(1 - n_{i', \tau'}^f)n_{i, \tau}^o n_{i, \tau'}^o, \quad (4.36)
\]

\[
P_{22} = \sum_{i \neq i'} \left( \prod_{\sigma, \tau, \tau'} \hat{n}_{i, \sigma}^f \right) n_{i', \tau}^o n_{i', \tau'}^o n_{i, \tau}^o n_{i, \tau'}^o, \quad (4.37)
\]

Here \(P_{00}\) projects out of the subband with two unoccupied localized sites, \(P_{02}\) with one unoccupied and one doubly occupied site, and \(P_{22}\) with two doubly occupied localized sites. The notations become more complex as more localized sites are considered in different energy excitations. To keep the derivation simple, we will map all of these projection operators to \(P_l\) defined for the Hubbard model, in which \(l\) represents the corresponding subband in the Anderson Model without energy degeneracy. When we consider a particle hopping between two energy subbands, it should be clear that the particle can only hop up or down to one of two energy subbands, \(U - \Delta\) or \(\Delta\), from the current energy level, even though there may be many other subbands closer in energy. Only the energy subband that the particle can reach will be considered in the perturbation and hence labelled with an index. These hoppings are called \(\text{inter-subband hoppings}\) as particles are transferring across an energy gap to different energy subbands. These kinds of hopping are believed to be responsible for the magnetic ordering of the lattice, since hopping between different singly occupied localized sites is spin-dependent, according to the Pauli exclusion principle.

On the other hand, particles can transfer within the same subband while the total amount of energy remains unchanged during the transfer. These hoppings are called \(\text{intra-subband hopping}\) and is most likely to impact on the conductivity of materials. This is due to the fact that charged particles can easily move around within the same band without crossing any energy gaps. They could even stop completely when no room is available for hopping without going across an energy gap to the next energy band. One of the assumptions of the canonical perturbation theory developed by CSO is that this intra-subband hopping is 'flat' \((ie.\, \text{independent of})\)
4. Perturbation Theory of the Atomic Limit

energy) and have no effect on the inter-subband hopping. Since the width of each subband is proportional to $V^2$, the theory basically requires that

$$\frac{V^2}{\Delta} \ll 1, \quad \frac{V^2}{U - \Delta} \ll 1$$  \hspace{1cm} (4.38)

We therefore assume these conditions to be satisfied in the following derivations.

The Hamiltonian for the periodic Anderson model is already given by Eqs. (4.1) and (4.2) in real space. All electron operators are defined as usual, but for the high temperature superconducting oxides (HTSO) they can be re-interpreted as holes. In particular, the operators $f_i^\dagger$ and $c_j^\dagger$ can be treated as creating holes in a $d_{x^2-y^2}$ orbital on Cu$^{2+}$ site $i$ and in a $p_{x(y)}$ orbital on O site $j$, respectively. The different orbital hopping $V_{ij}$ and the transfer matrix in the conduction band $t_{ij}$ and the on-site Coulomb repulsion $U$ between the localized electrons however remain the same.

Without the hybridization or orbital coupling term from $H_V$, the localized particles remain at the same energy level forever since $H_0$ does not allow any particles in the f-orbital to move. $H_0$ therefore corresponds to intra-subband hopping and can be written as

$$H_0 = \sum_{l=0}^{\infty} P_l H P_l$$  \hspace{1cm} (4.39)

The hybridization term $H_V$ allows a particle in a localized site to hop onto the conduction band and then arrive at another localized site. This kind of hopping involves changes in the occupation of localized sites which in turn changes the energy content of the whole system. In some cases, the hopping would also depend on the spin of the hopping particle, if the localized sites involved is already singly occupied, as double occupation of two particles of the same spin is prohibited. $H_V$ therefore describes the inter-subband hopping.

$$H_V = \sum_{l \neq l'}^{\infty} P_{l'} H P_l.$$  \hspace{1cm} (4.40)

In the following calculation, $H_V$ is treated as perturbation and the unperturbed Hamiltonian $H_0$ is solved as a two band model in the atomic limit (ie. $V = 0$).

Following the work of CSO, we first assume that the eigenenergies of $H_0$ consists of disjointed regions in each of which the spectrum can be either continuous or discrete. The widths of the regions are also assumed to be much smaller than the separations between them. This is shown in figure 4.6 with thick lines representing the width of each subband, and the large spacing as the separation between them. Another feature of this representation is that the perturbation $H_V$ has zero matrix elements between the eigenstates in the same subband. That is $P_l H V P_l = 0$. 
4.2. Canonical Perturbation Theory

We begin with the canonical transformation as follows:

\[ \tilde{H}(\varepsilon) = e^{-i\varepsilon S}(H_0 + \varepsilon H_V) e^{i\varepsilon S} \]  

(4.41)

where \( \varepsilon \) is formally treated as small except at the end of the calculation when we set it to one. Expanding the transformation as described in chapter 5,

\[ \tilde{H}(\varepsilon) = H_0 + \varepsilon (H_V - i[S, H_0]) + \sum_{n=2}^{\infty} \frac{(-i\varepsilon)^n}{n!} ([S, H_0]_n + i n ([S, H_V]_{n-1}). \]  

(4.42)

The parameter \( S \) will be determined by the condition that

\[ H_V - i[S, H_0] = 0, \]  

(4.43)

as in chapter 5. This eliminates the first order terms of all intra-subband hopping, and simplifies the Hamiltonian to

\[ \tilde{H}(\varepsilon) = H_0 + \sum_{n=2}^{\infty} \frac{(n-1)(-i)^{n-1} e^n}{n!} ([S, H_V]_{n-1}. \]  

(4.44)

The next step is to find the expression for \( S \) from the condition in Eq. (4.43). We substitute Eqs. (4.39) and (4.40) into Eq. (4.43) and apply the projection operators \( P_j \) and \( P_k \) to Eq. (4.43) from both sides to get

\[ P_j H P_k (1 - \delta_{jk}) + i P_j H P_j (P_j SP_k) - i (P_j SP_k) P_k H P_k = 0. \]  

(4.45)

When \( j \neq k \), it simplifies to

\[ P_j S P_k = \frac{P_j H P_k}{E_j - E_k} \]  

(4.46)

in which the intra-subband hopping terms \( P_j H P_j \) and \( P_k H P_k \) are approximated by their proper expectation values of the energies, \( E_j = P_j H_0 P_j \) and \( E_k = P_k H_0 P_k \). It is obvious from the equation that these two energy states at \( j \) and \( k \) are coupled by the inter-subband hopping \( P_j H P_k \). As long as the assumption that inter-subband hopping is much greater in energy than intra-subband hopping holds, the approximation of the two energy states at two different levels \( i \) and \( j \) is justified.

When \( j = k \), Eq. (4.45) can be rewritten as

\[ [P_j H_0 P_j, P_j SP_j] = 0, \]  

(4.47)

to which the general solution can be found from Ref. [75] as

\[ P_j S P_j = \sum_i P_i z P_i \]  

(4.48)
where $z$ is an arbitrary operator. To determine $S$, we observe from Eq. (4.47) that $S$ is independent of $H_V$ and so within the subband, we can assume $H_V = 0$ or $P_j SP_k = 0$. Therefore,

$$
S = \sum_{j,k} P_j SP_k = \sum_j P_j SP_j = \sum_j P_j ZP_j,
$$

(4.49)

where $Z$ is an arbitrary operator. The unperturbed Hamiltonian is transformed to

$$
\tilde{H}(\varepsilon = 1) = e^{-i\sum_j P_j ZP_j H_0} e^{i\sum_j P_j ZP_j} = H_0
$$

(4.50)

as $P_j ZP_j$ commutes with $P_j H_0 P_j$. The eigenstates of the Hamiltonian $|\Phi_j\rangle$ for a different energy level $j$ will also transform as

$$
|\tilde{\Phi}_j\rangle = e^{-i\sum_k P_k ZP_k} |\Phi_j\rangle = e^{-i\xi_j} |\Phi_j\rangle,
$$

(4.51)

where $\xi_j$ is the eigenvalue of $P_j ZP_j$ with respect to the state $|\Phi_j\rangle$ which takes advantage of the commutation of $P_j ZP_j$ with $P_j H_0 P_j$. The effect of the $P_j SP_j$ terms in the canonical transformation simply adds a phase factor to each eigenstate of the basis function. Since the basis can be uniquely defined only within a phase factor, Hamiltonian becomes

$$
\tilde{H}(\varepsilon) = H_0 - \sum_{n=2}^{\infty} \frac{(n-1)(i\varepsilon)^n}{n!} \sum_{all\ klm} \left( \sum_{j=0}^{n} \frac{(-1)^j n!}{j!(n-j)!} E_{k_{j+1}} \right) P_k SP_l S \cdots SP_{k_{n+1}}.
$$

(4.52)

### 4.2.1 The coefficient of inter-subband hopping

Despite the complicated expression shown in Eq. (4.52), the result is in fact quite simple. In the following, we will show how the terms can be calculated from diagrams and what their physical significances are. To our knowledge, this approach has not been taken before and the calculation in the past for each coefficient relied upon tedious term by term expansion. In the following, we will present a new diagrammatic approach and show with examples how it works for any order.

In the first step, we re-write the result in Eq. (4.52) to a slightly different form, as suggested by CSO, using

$$
\sum_{j=0}^{n} \frac{(-1)^j n!}{j!(n-j)! E_{k_{j+1}}} = \sum_{j=1}^{n} \frac{(-1)^j (n-1)!}{(j-1)!(n-j)!} (E_{k_{j+1}} - E_{k_j}),
$$

(4.53)

The difference of two adjacent indices $k_{j+1}$ and $k_j$ must be either $+1$ or $-1$ for one of the two possible hoppings, to either the upper subband or the lower subband. Eq.
4.2. Canonical Perturbation Theory

(4.53) turns the absolute energy scale into a relative scale and as a result, only the change in energy is needed when a charge carrying particle hops between different energy subbands. The next step comes from the substitution of $S$ by $H_V$ in Eq. (4.46). This step removes the dependence of the result from the form of $S$, and provides a better understanding of the physics in terms of the Hamiltonian. Using Eq. (4.46), we can write $P_k_1SP_k_2S\cdots SP_{k+n}$ as

$$(-1)^nP_k_1H_VP_k_2H_V\cdots H_VP_{k+n}$$

in which we used $H_V$ instead of $H$ in the expression since we know $H_0$ alone is responsible for the intra-subband hopping and has no effect to the hopping between subbands. Similarly as we mentioned before, $H_V$ is responsible for the hopping between subbands and has no effect on the intra-subband hopping. In other words, we have always assumed that

$$P_iH_0P_j = 0, \quad P_iH_VP_i = 0, \quad \text{all for } i$$

which separates the Hamiltonian into diagonal and off-diagonal parts. After using these two simplifications, the result Eq. (4.52) becomes

$$\hat{H}(\varepsilon = 1) = H_0 + \sum_{n=2}^{\infty} \frac{n-1}{n!} \sum_{all \ k_m} \sum_{j=1}^{n} (-1)^{j-1}C_{j-1}^n \frac{E_{k_{j+1}} - E_{k_j}}{\prod_{j=1}^{n}(E_{k_{j+1}} - E_{k_j})}.$$  (4.56)

where $C^n_j$ is the number of ways to arrange $j$ identical particles in $n$ identical places and $n$ represents the total number of hoppings between the corresponding subbands in the summation. It has the familiar binary form of the coefficients of $(x + y)^n$ expansion. This makes the calculation for the coefficients of the second term on the right hand side of Eq. (4.56) as simple as drawing diagrams and summing numbers, using the following steps.

1) Draw a diagram to show the hopping of electrons across different subbands as shown in figure 4.7 for example. In each unit of the horizontal axis, only one hopping is allowed to either one of the two subbands $U - \Delta$ and $\Delta$ from the current band. Since we are only interested in the eigenstates of the system, the final hopping must return to the original band. This ensures the number of hops is even and so is the value of $n$. Hopping to the upper subband is carried out by an electron and an 'e' label is written with it, whereas hopping down to the lower subband is done by a hole and so a 'h' is used instead.
Table 4.1: The coefficients of binary series with order $n - 1$ where $n$ is the number of inter-subband hops.

<table>
<thead>
<tr>
<th>$n - 1$</th>
<th>$C_j^{n-1}$ for $j = 1 \ldots n - 1$,</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1 2 1</td>
</tr>
<tr>
<td>3</td>
<td>1 3 3 1</td>
</tr>
<tr>
<td>4</td>
<td>1 4 6 4 1</td>
</tr>
<tr>
<td>5</td>
<td>1 5 10 10 5 1</td>
</tr>
<tr>
<td>6</td>
<td>1 6 15 20 15 6 1</td>
</tr>
<tr>
<td>7</td>
<td>1 7 21 35 35 21 7 1</td>
</tr>
<tr>
<td>8</td>
<td>1 8 28 56 70 56 28 8 1</td>
</tr>
<tr>
<td>9</td>
<td>1 9 36 84 126 126 84 36 9 1</td>
</tr>
</tbody>
</table>

II) Write down the binary series of order $n - 1$ from table 4.1 under the arrows in the diagram in exactly the same order as shown in the table, where $n$ is the total number of hoppings. In this way, there should be a number associated with each hop in the diagram. We then multiply a $-1$ to every second term of the selected series. For example, when $n = 4$ with four hops allowed in the transition, the binary sequence should be $1 - 3 3 - 1$. Note that the number of terms in the series is always even when $n$ is even.

III) Multiply each number of the binary sequence with the energy difference due to the associated hops. If the participant of a hopping is a hole, the energy change due to this hopping will attract a negative sign.

IV) Sum the series up, multiply the sum with a pre-factor of $(n - 1)/n!$ and then divide it by the product of each energy change during hopping.

As only two subbands are available to hop to from any current energy band, we use the symbol $\Gamma$ to represent the energy level difference of the upper band, similar to the use of symbol $\Delta$ for the lower band. In the Anderson model, we have $\Gamma = U - \Delta$ whereas in the Hubbard model $\Gamma = \Delta = U$.

For second order scattering, the result of the summation in Eq. (4.52) contains $1/\Delta$ and $1/\Gamma$ for a hole hopping down and returning or an electron hopping up and returning to the ground state energy level. Figures 4.8, 4.9 and 4.10 show three more complicated examples of four hopping processes. The coefficients calculated for all four examples are listed below in the same order as the figures 4.7, 4.8, 4.9.
4.2. Canonical Perturbation Theory

and 4.10 are shown.

\[
\frac{1}{8} \cdot \frac{4(\Delta + \Gamma)}{\Delta^2 \Gamma^2}, \quad \frac{1}{8} \cdot \frac{8\Delta}{\Delta^4}, \quad \frac{1}{8} \cdot \frac{-4\Delta}{\Delta^4}, \quad \frac{1}{8} \cdot \frac{2(\Delta - 3\Gamma)}{\Delta^2 \Gamma^2}
\]

Some features of the coefficients become obvious from the diagram. Firstly, due to the symmetric nature of the binary series, the coefficient calculated from the diagram is the same except for an extra factor of \((-1)^{n+1}\), if we run the hopping back in time. That is to say if we look at the diagrams from the right to left and revert all the hopping arrows. The results are the same for odd orders (when \(n\) is odd), whereas for even orders they differ by a minus sign. Secondly, the coefficients are completely different if the hopping sequence starts from a different energy subband, because the hopping pattern depends on the initial condition. In other words, for the same term in the Hamiltonian, the coefficient changes if the model changes from quarter filling to half filling, or three-quarter filling.

Thirdly, it is possible, and in some cases quite easy, to obtain a coefficient for any value of \(n\). This could lead to the evaluation of the dominant coefficient and hence the determination of the basic characteristics of the model quickly. For example, we can calculate the hopping of example two in figure 4.8 to any order quite easily. As in step 3 of the above method, the change of sign in the energy difference during hopping compliments the change of sign in the binary series. The coefficient calculated from the sum of the products for \(n^{th}\) order hopping is therefore,

\[
\frac{n-1}{n!} \sum_{j=1}^{n} -C_{j-1}^{n-1} \frac{1}{\Delta^{n-1}(-1)^{[(n+1)/2]}} = (-1)^{[(n-1)/2]} \frac{n-1}{n!\Delta^{n-1}2^{n-1}},
\]

where \([(n+1)/2]\) is assumed to be the integer part of the division. For even values
of $n$, it merely is $n/2$; but for odd values of $n$, it becomes $(n+1)/2$. Also, the sum of a binary series $\sum_{j=0}^{n} C_j^n = 2^n$ is used, to get the final result. Since the sign of this coefficient changes for every second even order or every second odd order, it is not possible to tell whether the system could gain energy from this kind of hopping. A similar result is also found for hopping between $P_1$ and $P_2$ states, with the sign reversed and the energy difference changed to $\Gamma$.

4.2.2 The inter-subband hopping

The next task is to understand the last factor of the second term in Eq. (4.56) which contains the projection operator $P_0, P_1$ and $P_2$ projecting the perturbed Hamiltonian $H_V$ onto different energy subbands. This factor determines the interaction of particles and gives rise to the magnetic properties of the system. We will start with a simple case of a single scattering process and build up more complicated processes based on this.

During a single scattering process, such as $P_1 H_V P_0$, the process turns an impurity particle to a conduction carrier, or vice versa. The detail of this process in this
Figure 4.11: The first order hybridization diagram from a localized electron to a conduction electron. The curved line shows the propagation of a conduction electron, whereas the horizontal line shows the propagation of an impurity electron. This is one of the two virtual processes.

Figure 4.12: The first order hybridization diagram from a conduction electron to a localized electron. The lines are defined the same as in figure 4.11. Similar to figure 4.11, this is the second virtual process.

Example is

\[
P_1HVP_0 = \prod_{j',\sigma'} \hat{n}_{j',\sigma'} f_{i,j,\sigma} c_{j,\sigma} \sum_{k} (1 - n_{k,i}^f)(1 - n_{k,i}^c) \prod_{l \neq k,\sigma''} \hat{n}_{l,\sigma''}^f,
\]

which destroys a particle in the conduction band at site \( j \) and turns it into a particle at the localized site \( i \) of spin \( \sigma \). The other process which hybridizes a particle from a localized site to the conduction band is,

\[
P_1HVP_2 = \prod_{j',\sigma'} \hat{n}_{j',\sigma'} V_{i,j} f_{i,\sigma} c_{j,\sigma} \sum_{k} n_{k,i}^c n_{k,i}^f \prod_{l \neq k,\sigma''} \hat{n}_{l,\sigma''}^f,
\]

\[= V_{i,j} c_{i,\sigma} f_{j,\sigma} n_{i,-\sigma} \prod_{i \neq j,\sigma'} \hat{n}_{i,\sigma'}^f \]

(4.59)

Both of these processes can be drawn as diagrams in figure 4.11 and 4.12 which involve the hybridization of only one particle. These simple diagrams will serve as building blocks for our more complicated processes. They are sometimes referred to as virtual processes as neither of the two can happen by themselves alone in a real process. We would have otherwise ended up with an unphysical model in which only one kind of particle would populate the whole system. Also neither of these two processes can return the system to its original state. In other words, they are not eigenstates and hence cannot have any impact on the ground state of the system, purely by themselves.

The picture is quite different if we consider combinations of these two processes. Let us first consider the second order term that contains two of these scattering
Perturbation theory of the Atomic Limit

Figure 4.13: All four second order scattering diagram. The lines and symbols are the same as defined in figure 4.11.

processes. One example of these processes in terms of projection operators is

\[ P_1 H_V P_0 H_V P_1 = (P_1 H_V P_0) (P_0 H_V P_1) = (P_1 H_V P_0) (P_1 H_V P_0)^\dagger, \]

which makes use of the fact that \( H_V = H_V^\dagger \), \( P_n = P_n P_n \) and \( P_n^\dagger = P_n \) for \( n = 0, 1 \) or 2. Substituting the particle operators into it, one gets

\[ P_1 H_V P_0 H_V P_1 = \prod_{i \neq i', \sigma'} \hat{n}_{i', \sigma'}^f V_{i,i'} (f_{i,i',-\sigma}^f c_{i,\sigma} c_{i',\sigma,-\epsilon} + c_{i,\sigma} c_{i',\sigma}^\dagger \hat{n}_{i,\sigma}^f) \]

We have used the fact that the projection operators are orthogonal to obtain a single index for the f-electron operator. In fact, it indicates that two scattering processes must involve the same impurity site for the same site index, to generate a \( P_0 \) eigenstate in the scattering process. It would have otherwise been two uncorrelated virtual processes that would not result in an definite unoccupied impurity site. This can also be drawn as diagram (a) and (b) shown in figure 4.13 in which the \( c_{j'} \) particle is actually created before the \( c_j \) particle is destroyed. When \( j \neq j' \), we can interpret it the other way around, that the \( c_j \) particle is destroyed and has turned into a \( c_{j'} \) particle, with an extra minus sign. However, when \( j = j' \) an additional term is needed to satisfy the commutation relation \( c_{j,\sigma} c_{j,\sigma}^\dagger = 1 - c_{j,\sigma} c_{j,\sigma} \) which could nevertheless be renormalized into the unperturbed Hamiltonian. Strictly speaking, this process should be written as

\[ \prod_{i \neq i', \sigma'} \hat{n}_{i', \sigma'}^f V_{i,i'} (f_{i,i',-\sigma}^f c_{i,\sigma} c_{i',\sigma,-\epsilon} - c_{i,\sigma} c_{i',\sigma}^\dagger \hat{n}_{i,\sigma}^f + \delta_{j,j'} \hat{n}_{i,\sigma}^f) \]

The extra unity term from the commutation relation will have a profound effect to the final result of the interactions.

For the process with a doubly occupied intermediate state, we have

\[ P_1 H_V P_2 H_V P_1 = \prod_{i \neq i', \sigma'} \hat{n}_{i', \sigma'}^f V_{i,i'} (c_{i',\sigma} c_{j,\sigma} c_{i,\sigma}^\dagger f_{i,i',-\sigma}^f + c_{i,\sigma} c_{i',\sigma}^\dagger \hat{n}_{i,\sigma}^f) \]
and the corresponding diagrams shown in diagram (c) and (d) of figure 4.13. Unlike the former process, the operators in this case appear in the appropriate order and therefore the unity term that appears in Eq. (4.63) is absent. In addition, the spins of the impurity site in the last terms of both Eq. (4.62) and Eq. (4.64) are opposite which implies both processes corresponding to these two terms can take place at the same time, due to the orthogonality of the up and the down spin particles. Each of these two processes can also be written in terms of the product of the c-particle and the localized particle spin operators, using the spin operator transformation,

\[ S_i^+ = f_{i,i}^\dagger f_{i,i}, \quad S_i^- = f_{i,i}^\dagger f_{i,i}, \quad S_i^z = \frac{1}{2}(n_{i,i}^f - n_{i,i}^f), \]
\[ s_{ij}^+ = c_{i,i}^\dagger c_{j,i}, \quad s_{ij}^- = c_{i,i}^\dagger c_{j,i}, \quad s_{ij}^z = \frac{1}{2}(c_{i,i}^\dagger c_{j,i}^\dagger - c_{i,i}^\dagger c_{j,i}). \]  

This turns the sum of the two processes multiplied with their corresponding coefficients \(-1/\Delta\) and \(1/\Gamma\) into

\[ -\sum_{i,j} \frac{V_{ij}V_{ji}}{\Delta} + \sum_{i,j,j'} \left( \frac{1}{\Delta} + \frac{1}{\Gamma} \right) V_{ij}V_{jj'} S_i^t S_j^{t'} + \sum_{i,j,j',\sigma} \frac{1}{2} \left( \frac{1}{\Delta} - \frac{1}{\Gamma} \right) V_{ij}V_{jj'} c_{i,j,\sigma} c_{j',\sigma}. \]  

This result agrees with the result from Zaanen and Oleš [97] and the result from Schrieffer and Wolff [78] using canonical transformation in chapter 5.

In the third order process when we have three of the elementary diagrams connected together, a c-electron or f-electron would turn into each other as in the first order virtual processes. It is shown explicitly in figure 4.14 that they indeed are virtual processes themselves. As explained previously, they would not contribute to the eigenstate of the Hamiltonian.

However, adding one more elementary process to the third order processes results in four different possible processes. The first one involves particle jumping down to the lowest subband \(P_0\) and returning back to the conduction band two times in a row. The second has one going down to the lowest state \(P_0\) and then followed by another one going up to the next highest state \(P_2\). The third process is the reverse
of the second process while the fourth one contains particles going up to the highest level $P_2$ and returning two times in a row.

We do not show the equations here, as the complexity is likely to mislead the actual physical insight of the processes. The corresponding diagrams for one of the two spin configurations are however displayed in figure 4.15. The fourth order processes can in fact be considered as two second order processes occurring one after another. The created c-particle from the first process does not need to be the same c-particle destroyed in the second process, since the projection operator is indiscriminatory between c-particles. Zaanen and Oleś [97] use this property to find the carrier to carrier interaction mediated with a localized particle.

All fourth order processes resemble the second order processes, if the c-particle keeps scattering the same impurity site. In fact, this is the starting point of the 'poor-man's scaling' when Anderson et al [7] considered the middle c-particles as virtual scattering at the upper and the lower band edge of the conduction band in momentum space. After eliminating this virtual scattering processes by summing the indices $j''$, $j'''$, they obtained a result that is very similar to the one he started with. The difference of these two results therefore establishes a flow equation which then allows them to find the fixed point in the phase space of the model and define the boundary between the ferromagnetic and the anti-ferromagnetic phase.

In our case, the c-particle is represented in real space and cannot be considered as virtual. Also, the impurity site is unique in Anderson's poor-man's scaling for the single impurity model, whereas the c-particle can scatter particles at any impurity
4.2. Canonical Perturbation Theory

site of a lattice. Nevertheless, our model does provide the real space representation that allows us to investigate the motion of particles in a lattice.

Furthermore, only two out of four processes in figure 4.15 retain the spin configuration between the initial and the final stages. The other two processes have the spins reversed in the final outcome. Only the first two would therefore be considered as the eigenstate of the process in this order.

4.2.3 Some interesting results

One could work out the higher order results by putting different combinations of the processes from the previous order, as was done from the second to the fourth order above. The physics is indeed very much the same. For example, only half of the processes would have the final spin configuration the same as the initial stage. The number of processes would increase exponentially in the power of two. Every odd order process would correspond to the hybridization exchange which does not contribute to the eigenstate of the Hamiltonian. Out of all these repetitive processes, it is however rather interesting to pick out those that could give rise to some interesting phenomena.

The first one we are going to consider is when the carrier c-particles keep scattering back to the same impurity site. Zaanen and Oleś investigated this case only to the fourth order [97], but it is indeed possible to calculate it to infinite order. The first step is to pick the processes that have the same initial and final states. As mentioned before, half of the processes out of each order satisfy this condition. The second step is to find out the coefficients corresponding to each of these processes. It seems a difficult task at first, but looking back at the calculation of the coefficients, it is not difficult to see that the largest contribution must come from the particle hopping which compliments the sign change of the binary series, as mentioned previously\(^1\). Only two such processes exist which are shown in figure 4.16 and 4.17. The sum of each product in both of these cases attains the extreme value.

The final step is to find the coefficients of these two processes for the \(n\)th order. To simplify the calculation, we define the second order process with two symbols \(\bigtriangledown\) and \(\Delta\). \(\bigtriangledown\) represents Eq. (4.63) without the unity factor and \(\Delta\) represents Eq. (4.64). These symbols are chosen simply because of their resemblance to the corresponding hopping energy diagram. The \(n\)th order process for \(\Delta\) can therefore

\(^1\)Either \(\Gamma\) or \(\Delta\) contributes to the first excitation energy, depending on their relative magnitude. It is therefore safe to consider both of them as the most significant processes and all other processes as a result of some combination of them would be the second order effect.
be written as

$$V^n \Delta \Delta \cdots \Delta$$

(4.68)

and the corresponding coefficient is

$$\frac{n - 1}{n!} \frac{(-1)^{n/2} 2^{n-1}}{\Gamma^{n-1}}$$

(4.69)

in which $n$ is always even. We have also assumed that the hybridization factor $V$ is independent of the site index for simplicity. For the second extreme process, the equation in terms of our defined symbols is

$$V^n (1 + \nabla)(1 + \nabla) \cdots (1 + \nabla)$$

(4.70)

or

$$V^n (1 + n/2 \nabla + C_2^{n/2} \nabla^2 + \cdots + +n/2 \nabla^{n/2-1} + \nabla^{n/2})$$

(4.71)

The coefficient corresponding to this process is

$$\frac{n - 1}{n!} \frac{(-1)^{n/2} 2^{n-1}}{\Delta^{n-1}}$$

(4.72)

It is obvious from this equation that lower order processes can actually be generated from higher order processes, due to the unity factor in the second order scattering terms. All terms must therefore be considered when the $V$ is not small. Now, we can sum the coefficients up to the $n$th order. The first one is the unity term which arises from multiplying all the ones together from the equation. This is also the only unity factor possible from the $n$th order process. After multiplying the energy
4.2. Canonical Perturbation Theory

factor, the term becomes

$$\sum_{i=1}^{n/2} \frac{(-1)^i(2i - 1)(2^{2i-1})}{(2i)!\Delta^{2i-1}} V^{2i},$$

and then summing up to infinite order for $n = 2i \to \infty$, we get

$$V \sin\left(\frac{2V}{\Delta}\right) + \frac{\Delta}{2}\left(\cos\left(\frac{2V}{\Delta}\right) - 1\right).$$

It is rather surprising that the summation of infinite order terms reduces to a simple trigonometric equation. This term nonetheless only renormalizes the zero-point energy of the system in the Hamiltonian. The impact is entirely different with the first order process from which the sum to $n$th order process is given by,

$$\frac{1}{2} \sum_{i=2}^{n/2} \frac{(-1)^i(2^{2i-1})}{(2i - 2)!\Delta^{2i-1}} V^{2i}$$

and summing $n$ up to infinity yields

$$\frac{V^2}{\Delta} \left(\cos\left(\frac{2V}{\Delta}\right) - 1\right).$$

This coefficient is in general two order of magnitude in $V$ smaller than the zeroth order term and so may be neglected when considering the dominating characteristic of the model. All other terms from the result of the $n$th order process can be calculated in a similar fashion, but their significances decrease as the order of the term increases.

Interestingly, the highest order term of the $n$th order process has exactly the same coefficient as the unity factor, from the last term of Eq. (4.71). This factor, when combined with the $n$th order $\Delta$ process, produces the spin interaction to $n$th order. Let us start with the second process in Eq. (4.68) that contains the symbols $\Delta$. The multiplication of the coefficient in Eq. (4.69) with the process to the $n$th order is

$$\sum_{i=1}^{n/2} \frac{(-1)^i(2i - 1)(2^{2i-1})}{(2i)!\Gamma^{2i-1}} V^{2i},$$

and the corresponding sum to infinite order is

$$V \sin\left(\frac{2V}{\Gamma}\right) + \frac{\Gamma}{2}\left(\cos\left(\frac{2V}{\Gamma}\right) - 1\right)$$

The sum of this process with the other process in Eq. (4.74) gives rise to the spin interaction between the localized and the carrier particles, as was shown in
the second order process. Therefore, we can assemble from these two terms the following interaction in the Hamiltonian

$$V(F(\frac{2V}{\Delta}) + F(\frac{2V}{\Gamma}))S_{j_1}^c s^c_{k_1} S_{j_2}^c s^c_{k_2} \cdots S^c_{j_{\infty}} s^c_{k_{\infty}}$$  \hspace{1cm} (4.79)$$

in which $F(x) = \sin(x) + (\cos(x) - 1)/x$. The strength of this interaction therefore varies with the upper and the lower energy subbands alone. One distinct feature of the trigonometric function in the above expression is that it oscillates between positive and negative values as $V$ varies. In other words, ferromagnetic coupling could be energetically favourable if the coefficient of the spin interaction becomes negative at some value of $V$. This value of $V$ must however be small for the perturbation to remain convergent. A similar equation will appear in chapter 6 when we consider the spin interaction of the lattice model using canonical transformation. No restriction on $V$ is required in that case when the model is transformed canonically. More detail will be discussed in that chapter.

Another case of interest is when the particle interactions are limited to the nearest neighbours. This significantly reduces the number of terms in the canonical interaction.
p perturbation expansion and provides a more realistic picture of real materials. In this case, two localized particles at two different locations \( m \) sites apart can interact with each other through the \( m \)th order process or higher, mediated with carrier particles. This \( m \)th order process can be calculated in a similar fashion with the recursive scattering of a single impurity in a lattice case. The difference is only that the carrier particles in this case are correlated which provides a link between two impurity particles of \( m \) sites apart. The interaction between the impurities however is only through nearest neighbour and so two impurities of \( m \) sites apart would need at least a \( m \)th order process to interact with each other. The largest contribution of all processes for spin interaction \( J \) of particles \( m \) sites apart is the \( m \)th order term of Eq. (4.79). The coefficient in Eq. (4.79) for finite order summation can be considered as the sum of two functions, one from Eq. (4.73) and another one from the same Eq. (4.73) with \( \Delta \) replaced by \( \Gamma \). Using the same plot of Eq. (4.73) in figure 4.18, one can then make two plots for the two corresponding functions by re-scaling the x-axis according to the value of \( \Delta \) and \( \Gamma \) respectively. The final coefficient of the spin interaction for finite order summation is given by the sum of these two plots. Figure 4.18 shows six different cases in which the curve named “nearest neighbour” corresponds to the coefficient \( J \) of two nearest neighbour impurities. The curve labelled “next nearest neighbour” shows the spin interaction between two impurities of one lattice site apart. The rest of the curves are named in the same fashion. Except for the case with ‘nearest neighbour’, it is obvious that all other curves are either positive definite or negative definite, depending on the starting order of summation. This means that once the relative distance of the two impurities is determined, the sign of the spin coefficient is fixed, regardless of the values of \( \Delta \) and \( \Gamma \). In this case, it is easy to see that the spin interaction is generally anti-ferromagnetic with the nearest neighbours, but ferromagnetic with the next nearest neighbours. This interaction becomes anti-ferromagnetic again with impurities three sites apart and ferromagnetic with impurities four sites apart. This seems to fit with the overall picture of an anti-ferromagnetic ground state, except for the possibility of ferromagnetic coupling between nearest neighbours for some values of \( \Delta \) and \( \Gamma \).

The discussion so far with the Anderson lattice model is only limited to the terms in the Hamiltonian, after canonical perturbation expansion. Although the coefficient of each term provides a great deal of information regarding the spin correlation and particle hopping, the thermodynamic properties of the model is missing from this analysis which is more likely to be revealed from the partition functions. Following the approach of Haldane for the single impurity model [41],
we will extend the method and apply it to the lattice model. The method was originally derived in the most general sense which makes the lattice case a natural extension of the original theory.

4.3 Thermodynamic properties

The motivation of Haldane to investigate the thermodynamic properties of the Anderson Model stems from the fact that two perturbative approaches of the same single impurity Anderson model appeared to contradict each other with respect to the magnetic susceptibility. The first approach was carried out by Scalapino [77], Keiter and Kimball [53] who investigated the single impurity Anderson model using perturbation theory at the atomic limit and found leading logarithmic terms of the same form as those occurring in perturbation expansions for the Kondo model. Following the approach of the Kondo model, they assumed that these logarithmic terms must be controlled by a high-energy cut-off which they associated with the conduction electron bandwidth. However from the perturbation expansion study of Yamada [95] for the interacting single impurity Anderson model and the numerical renormalization group approach of Krishna-murthy et al [57], the infinite bandwidth limit of the model was found to be well behaved, and the Kondo effect as proposed by Scalapino et al was absent. This inconsistency as concluded by Haldane [41] must be related to the terms neglected in the perturbation expansion by Scalapino which provide the cancellation of the conduction bandwidth dependence at high energies. More details of the argument can be found in his original paper [41].

To resolve this conflict, Haldane reformulated the perturbation theory at the atomic limit of the infinite bandwidth Anderson model, in terms of a sum of distinct 'histories' of the system on a circular 'imaginary-time' interval. This allows the model to be expanded in various limits with both logarithmic and non-logarithmic terms. His approach was formulated in such a generic way that it can readily be applied to the Anderson lattice model, as well as the single impurity model. In the following, we will follow his approach in the context of a single impurity model first and then extend the results to our lattice case, with the calculation of the magnetic susceptibility of the lattice model as an example. This approach has never been taken before in the literature. The d-orbital in Haldane's approach is replaced by our f-orbital formulation for consistency.
4.3.1 Haldane’s perturbation expansion in single impurity model

The Hamiltonian of the single impurity Anderson model in Haldane’s approach is divided into three parts, different to the conventional method. Because of this, we provide the Hamiltonian of the single impurity model here.

\[ H = \sum_\sigma H^c_\sigma + H^f + \sum_\sigma H^sf_\sigma \quad (4.80) \]

in which each part is defined as

\[ H^c_\sigma = \sum_{i,j} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma}, \quad (4.81) \]

\[ H^f = \sum_\sigma (\epsilon_f + \frac{1}{2} \sigma H_f) n^f_\sigma + U n^f_i n^f_j, \quad (4.82) \]

\[ H^sf_\sigma = \sum_{i,j} V_{i,j} (f_{i,\sigma}^\dagger c_{j,\sigma} + \text{H.C.}) \quad (4.83) \]

The first part \( H^c_\sigma \) describes the conduction band of the model which by itself is the Hamiltonian of a free electron model. \( t_{i,j} \) defines the hopping of free electrons between nearest neighbours. The second part \( H^f \) is the Hamiltonian of electrons in the localized site. The localized site in our case is assumed to be the f-orbital, but is equally applicable to any other localized orbitals such as d-orbital. It contains the energy content of each localized electron and the Zeeman term which includes an external magnetic field related Hamiltonian \( H_f \). The Coulomb repulsion for double occupation is also included in this Hamiltonian. The last part \( H^sf_\sigma \) of the Hamiltonian is the hybridization which couples the carrier and the localized particles together, without which the model could have been diagonalized and solved exactly.

Similar to our previous approach, the Anderson model is subdivided into different energy levels depending on the configuration of the localized sites. For example, if the ground state contains purely singly occupied impurity sites, the next higher state in energy contains either one double occupation or one empty impurity site. Combinations of different kind of this electronic configurations at different sites in the lattice create many more states with different energy excitations. Since these energy states are orthogonal, we can rewrite the Hamiltonian in terms of these new energy states.

\[ H = \sum_{\alpha} E_{\alpha} X_{\alpha\alpha} + \sum_{\alpha,\beta} V_{\alpha\beta} X_{\alpha\beta} \quad (4.84) \]

where the basis operators \( X_{\alpha\beta} \) are defined by \( X_{\alpha\beta} = \langle \alpha | \beta \rangle \) and each state \( \{ \langle \alpha | \} \) represents the lattice of one electronic configuration at one particular energy level.
The first term which contains the diagonal elements corresponds to the electron interaction within the same energy band, whereas the second term containing the off-diagonal elements corresponds to inter-subband interactions. With the assumption that $V$ is small relative to other parameters, the second term can be handled as the perturbation to the first diagonalized Hamiltonian.

Using this form of Hamiltonian, we can re-write the partition function using the well-known identity.

$$Z(H^0 + H^1 + H^2 + \ldots) = \text{Tr}[\exp(-\beta(H^0 + H^1 + H^2 + \ldots))]$$

$$= \text{Tr}[e^{-\beta H^0} T(e^{-\beta H^1} x e^{-\beta H^2} x \ldots)] \ldots, (4.85)$$

where

$$H^1(x) = e^{x^2 H^0} x e^{-x^2 H^0}, \quad (4.86)$$

and $T$ is the $x$ ordering operator on the interval $[0,1]$, which orders strings of $x$ labelled Fermi or Bose operators with a sign convention appropriate to their commutation relations. Using the following $x$ dependency of the operators,

$$X_{\alpha\beta}(x) = X_{\alpha\beta} e^{x^2(E_{\alpha} - E_{\beta})}, \quad (4.87)$$

one obtains the representation

$$Z = \sum_{N} (-\beta)^N \sum_{\alpha_1 \ldots \alpha_N} \int_0^1 dx_N \int_0^{x_N} dx_{N-1} \int_0^{x_2} dx_1$$

$$(V_{\alpha_1} a^N V_{\alpha_N} a^N \ldots V_{\alpha_2} a^1 e^{-\beta \sum_\alpha E_\alpha T_\alpha \{x_i\}}), \quad (4.88)$$

where $T_\alpha \{x_i\}$ is the fraction of the interval $[0,1]$ spent in the state $|\alpha\rangle$.

$$T_\alpha \{x_i\} = \sum_{i=1}^N (x_i - x_{i-1}) \delta_{\alpha\alpha'}, \quad (4.89)$$

The physical meaning of $x_i$ can be interpreted as the history of the system with an event denoted by $i$. In other words, $x_i - x_{i-1}$ is the duration in history between two events, one happening at $x_i$ and another one at $x_{i-1}$. All of these events must be on a circle of unit length. Each ‘history’, as Haldane defined in [41], consists of intervals in which the system is in a given state $|\alpha\rangle$, separated by ‘events’ at $x_i$ where the state changes. The amplitude for a given history is a product of factors $e^{-\beta E_{\alpha^i}(x_i - x_{i-1})}$ for each interval between events $x_{i-1}$ and $x_i$ spent in the state $|\alpha^i\rangle$ and factors $-\beta V_{\alpha^i+1, \alpha^i}$ for each event.

By change of integration variables and use of relative intervals instead of absolute variables (ie. $(x_i - x_{i-1}) \rightarrow x_i$) as shown in figure 4.19, one can re-write Eq. (4.88)
4.3. Thermodynamic properties

Figure 4.19: An example of five different events occurring in the system on a unit circle. (a) shows the system starting from state $\alpha^1$ changes to state $\alpha^2$ due to an event happening at time $x_1$. The subsequent events at time $x_2, x_3, x_4$ and $x_5$ change the system from state $\alpha^2, \alpha^4, \alpha^5$ and back to $\alpha^1$. The time of the events can be anywhere on the unit circle, starting from zero at where the arrow points. The events in this picture are in the absolute time scale. (b) shows the same history of the system but the events are marked by relative time interval. $x_i$ now shows how long the system is in the corresponding state where $i = 1, 2, 3, 4, 5$.

\[
Z = \sum_{\text{distinct histories}} \prod_{i} (-\beta V_{\alpha^i, \alpha^{i-1}}) e^{-x_i \beta E_{\alpha^i}}. \tag{4.90}
\]

The details of the derivation are skipped here which can be found from the original paper [41]. The summation in the equation must be performed over distinct histories which contain different permutations of $n$ events on a unit circle. Some examples of histories are shown in figure 4.20. The formulation is indeed very general and could have been applied to other models such as the Hubbard model. The difference in our case is merely the number of energy states involved. With the single impurity Anderson model, only four different states are possible with zero, up-spin, down-spin and full occupation of the localized site. The history diagram for this is therefore very simple as only four different states are required to show the interactions. The four states are furthermore easily divided into two groups, one with empty and fully occupied impurity sites and the other with singly occupied up and down spin localized electrons. The first group can be associated with charge states while the second group as spin states. It is interesting to see that any hybridization between conduction electrons and localized electrons must change from a state of one group to a state of the other. For instance, if the hybridization involves the creation of a conduction electron, this electron must be destroyed at some later stage of the interaction to return the system to the original state. This creation and annihilation of conduction electrons therefore changes the state of the impurity site and causes the system to flip from one of the four possible states to another. Another way
4. Perturbation Theory of the Atomic Limit

Figure 4.20: Topologically distinct Anderson model diagrams up to second order. Internal broken lines are propagators of conduction electrons. The different states of the system due to the interaction of the conduction electrons are shown with symbols $\alpha, \beta$ and $\sigma$. These states can be from either the charge group or the spin group. Charge and spin states always occur in alternate fashion.

to view this is to connect every two events in the history diagram of the system which causes a state to change, with a broken line. This broken line represents the propagation of conduction electrons from one event to another, created and annihilated due to the hybridization process. This however does not necessarily mean that the annihilated electron is the one that was created in the same process. Since all electrons in the conduction band are indistinguishable and the conduction electron propagator is indiscriminatory among the conduction electrons, there is no restriction as to which electron is annihilated and which is created in any processes, provided the total spin of the system is conserved. With this in mind, the first three orders of the interaction between electrons in the conduction band and the impurity sites can be drawn as in figure 4.20.

The zero order diagram in (a) shows the system staying in the same state for the whole length of time. As only four states are possible, $\alpha$ can be any one of the four states. The first order diagram in (b) contains a conduction electron propagating from one side of the circle directly to the other and turns the state from a charge state to a spin state or vice versa. In other words, if $\alpha$ belongs to one of the two spin states in the figure, $\beta$ must be either one of the two charge states. The second order diagrams are slightly more complex as shown in (c) and (d). One of the possibilities in the second order diagram is that an electron created due to a change of state gets re-absorbed and turns the new state back to the original state, in the first half of its history (c), (d). A similar event happens again in the second half of the history with a different intermediate state. An alternative possibility contains two independent hybridizations happening one after the other in the history diagram (e). In this way
the four states that appear in the history are different. Since there are only four different possible states in the model, the two different states of the same group (e.g. up/down spin of the spin group or zero/two charge of the charge group) appear at the two ends of the circle. The third order diagrams (f.g.h,i,j) are all drawn based upon the same principle.

4.3.2 Magnetic susceptibility of the single and lattice impurity models

The history diagrams we developed in the last section can be used to write down the partition function. The full details of the derivation for the single impurity model can be found in Haldane’s paper [41] which we skip here. We will however use his results in this section and apply them to the lattice case. The main result of his paper is

$$\frac{Z_H}{Z_{He}} = \sum_{\text{distinct diagrams}} \left( \prod_i e^{-\beta E_{\alpha_i} x_i} \right) \left( \prod_j \frac{\beta \Delta}{\sin \pi |y_j|} \right),$$

where $\Delta(E) = \pi \sum_k |V_{kd}|^2 \delta(E - \epsilon_k)$. The amplitude for each distinct diagram consists of a factor $e^{-\beta E_{\alpha} x}$ for each interval of duration $x$ spent in state $|\alpha\rangle$ and a factor $\beta \Delta / \sin \pi |y|$ for each conduction electron propagator connecting events with a separation $y$. The measure of the sum over distinct diagrams is defined as

$$\sum_{\text{distinct diagrams}} = \sum_{\text{distinct sequences}} \frac{1}{p} \langle \pi \rangle \int_0^1 dx_i \delta(1 - \sum_i x_i),$$

where \(\{x_i\}\) are the durations of the intervals, and $p$ is the divisor of the sequence \((\alpha_1 \rightarrow \alpha_2 \rightarrow \ldots \alpha_N \rightarrow \alpha_1)\). $p$ is the number of identical sub-sequences the original sequence can be broken up into.

To make it easier to understand, we will illustrate the equation with the diagrams shown in figure 4.20 for the first three orders in the single impurity case. In the zero order case as shown in diagram (a), there is no electron propagation in the diagram as no interaction has taken place. This is the non-interacting term of the expansion and the partition function is therefore just:

$$Z_0 = e^{-\epsilon^-} + e^{-\epsilon^+} + e^{-h} + e^{+h},$$

in which we used the same notation as in Haldane’s paper [41] where $\epsilon^+ = \beta E_+ = \beta(E_d + U)$, $\epsilon^- = \beta E_- = -\beta E_d$ and $\sigma h = \beta E_\sigma = \frac{3}{2} \sigma \beta H_d$ for $\sigma = \pm 1$. The product term with $\sin \pi |y|$ in Eq. (4.91) is absent in the equation as no conduction electrons are involved in the diagram.
In the first order diagram, the starting state of the history \( \alpha \) can be either one of the charge states, for example. After emitting an free electron, the charge state turns into one of the two spin states. This free electron is subsequently re-captured and changes the spin state back to the original charge state. For each electron propagator shown as the broken line in the diagram, a factor \( \frac{\beta\Delta}{\sin \pi x} \) is multiplied by the state involved and so in this case, we have

\[
Z_1 = \beta\Delta \int_0^1 \frac{dx}{\sin \pi x} \left( e^{-\epsilon z} + e^{\epsilon z} \right) \left( e^{h(1-z)} + e^{-h(1-z)} \right)
\]

The second order diagrams are slightly more complicated but the rule is still the same. Figure 4.21 shows all possible configurations of the second order diagrams and the relationship of the integration variables. For instance if we start with one of the spin states \( \sigma \) at the top of diagram (a), an electron propagation turns it into a spin state which upon recapturing an electron of the same spin returns to the original state \( \sigma \). If this is the whole history of the system, it would have been a first order diagram. The system however undergoes the same process again, but this time the intermediate spin state could be different. The partition function based upon this is

\[
T_1 = (\beta\Delta)^2 \int_0^1 dx \int_0^1 dy \int_0^1 dz \frac{1}{2} \theta(1-x-y-z) \theta(z) \left( e^{-\epsilon^+z} + e^{-\epsilon^-z} \right) \left( e^{\epsilon^+y} + e^{\epsilon^-y} \right) (2 \cosh(h(1-x-y)))
\]

The factor \( \frac{1}{2} \) arises because each distinct diagram is generated twice by the sum over the spin and charge states. The \( \theta \) function is to ensure the relationship between the variables \( x, y \) and \( z \) are fulfilled as shown in diagram (d), i.e. all quadrants of the circle must be positive. Since we know \( x \) and \( y \) are between 0 and 1, it is not necessary to include \( \theta(x) \) and \( \theta(y) \) in the integrand. As for \( z \), the \( \theta \) function must be explicitly stated as the integrand is unchanged under exchange of the two intervals.
4.3. Thermodynamic properties

$z \leftrightarrow 1 - x - y - z$, and $(1 - x - y - z)$ could be negative if $z > 1 - x - y$. It is also possible to enforce that $1 - x - y - z > z$ arbitrarily so that the bottom quadrant is always larger than the top one. In this case, the factor $\frac{1}{2} \theta(1 - x - y - z)$ can be replaced by $\theta(1 - x - y - 2z)$ for all second order terms.

The conduction electron propagates in this diagram as far as the charge state, and so the distance for two electron propagations coincides with the two integration variables $x$ and $y$. The rest of the history belongs to the spin state $\sigma$ which is therefore of length $1 - x - y$.

The second diagram of the second order term is almost the same, except it starts with a charge state on the left of diagram (b) and finishes with the same charge state. The integration variables are also different for different states but the idea of summation is still the same. The partition function of this diagram can therefore be written as:

$$ T_2 = (\beta \Delta)^2 \int_0^1 dx \int_0^1 dy \int_0^1 dz \frac{1}{2} \theta(1 - x - y - z) \theta(z) \left( e^{-\epsilon^+(x+y)} + e^{-\epsilon^-(x+y)} \right) \cdot \left( 2 \cosh(h(1 - x - y)) + 2 \cosh(h(1 - x - y - 2z)) \right), $$

(4.96)

in which $x + y$ is the whole duration of the charge state but $z$ and $1 - x - y - z$ are the duration of the two spin states, as well as the two electron propagators. Note that $1 - x - y - z$ and $x + y + z$ are interchangeable in the electron propagation as they define the same chord distance between the same two points on a unit circle.

Using the same rules, it is not hard to find that diagram (c) in figure 4.21 can be written as

$$ T_3 = (\beta \Delta)^2 \int_0^1 dx \int_0^1 dy \int_0^1 dz \frac{1}{2} \theta(1 - x - y - z) \theta(z) \left( e^{-\epsilon^+x}e^{-\epsilon^+y} + e^{-\epsilon^-x}e^{-\epsilon^-y} \right) \cdot \left( 2 \cosh(h(1 - x - y - 2z)) \right) $$

(4.97)

and so the the second order partition function is

$$ Z_2 = T_1 + T_2 + T_3, $$

(4.98)

These are the results for the single impurity Anderson model. The simplicity of the result comes from the rather small number of possible states in the d-orbital of a single impurity model. With a lattice model, the number of possible states increases exponentially with the number of sites in the lattice. To the first three order terms, it is nonetheless still possible to obtain results for the lattice model using the same formulation. We first consider the zeroth order partition function with the corresponding pure circle diagram. As no interaction takes place in the
Perturbation Theory of the Atomic Limit

Figure 4.22: Periodic Anderson model diagrams of the zero (a) and first (b) order. The same notation as in figure 4.20 is used.

Figure 4.23: Periodic Anderson model diagrams of the second order. (a) and (b) are the two possibilities of the second order process. The same notation as in figure 4.20 is used.

zero order, the partition function is the $N$th power of the sum of the four different energy states in every impurity site where $N$ is the total number of impurity sites in the lattice.

$$Z_{0}^{PAM} = Z_{0}^{N} = (e^{-\epsilon} + e^{-\epsilon'} + e^{-\epsilon''} + e^{-\epsilon'''} )^{N}$$  (4.99)

This can also be described as the product of $N$ circles in figure 4.22, each one in one of the four possible states as in the single impurity model.

The first order term is again quite simple as only one electron propagation is allowed for all impurity sites. To make it even better, this electron in the first order term must propagate back to its original state to return the system to its eigenstate. This is shown in (b) of figure 4.22.

$$Z_{1}^{PAM} = N Z_{0}^{N-1} Z_{1},$$  (4.100)

The factor $N$ on the right hand side is to account for the $N$ possible ways this first order process can happen in $N$ lattice sites.

The second order term becomes a little more complex. One of the possibilities is when one of the $N$ impurity sites undergoes the second order process as if it was in the single impurity case. In this case both electron propagations happen at the same impurity site as shown in diagram (a) of figure 4.23. Since there are $N$ possible ways for this process to happen to any one of the $N$ impurity sites, a factor of $N$ is needed. The second possibility of the second order process is when two first order processes happen at the same time, shown in diagram (b) of figure 4.23, at two different sites in $N(N - 1)/2$ possible ways.

One might think that there could be a third possibility in which an electron propagates from site $i$ to site $j$ for example, and some time later another electron
propagates back from site \( j \) to \( i \) as shown in figure 4.24. This however has no difference to the second possibility we described above. As mentioned previously, the conduction electron propagation is indiscriminatory among the conduction electrons. It merely describes an electron being created at one point or annihilated at another without specifying which one in particular. It is therefore possible that electron \( c_1 \) is created in the conduction band, but electron \( c_2 \) is destroyed later in the process, provided the total spin of the system is conserved. In this respect, the broken lines in the unit circle of all the diagrams we mentioned above are merely to keep track of the spin of electrons being exchanged. These processes are therefore equivalent to two of the first order processes.

For the third order process, there are more possibilities as described in figure 4.25. The first possibility (a) has one site undergone the third order process, similar to the single impurity model. Another possibility is to have three first order processes taking place at different sites at the same time as shown in diagram (b). The last possibility contains one site undergoing a second order process while another site carrying out a first order process. The number of ways these processes can happen in a lattice of \( N \) sites is therefore \( N, C_2^N = n(n - 1)/2 \) and \( n(n - 1) \) respectively. The partition function can therefore be formulated in the same way as was done for the second order case.

Higher order processes are treated all in the same fashion and can be written down in a similar way. It is not difficult to see a general pattern starting to emerge after going through the first three orders. The results of the combinatorials suggest that they could have been generated from the \( N \)th power of the sum of the different single impurity diagrams. In fact, it is the nature of the Anderson model that electrons in the impurity sites do not interact directly with one another, even in the lattice case. They could only interact through the conduction electrons by means of hybridization. As the interaction between electrons in impurity sites and conduction band can be contained within a single history diagram\(^1\), the partition function of a \( N \) site lattice is the product of the sum of all possible diagrams of a single site. Since the \( N \) sites are essentially indistinguishable from one another, the partition

\[ \text{Figure 4.24: Two equivalent second order diagrams. The same notation as in figure 4.20 is used.} \]

\(^1\)As described before provided the total spin is conserved.
function becomes the $N$th power of the sum of all diagrams from the single impurity model. We can therefore re-use the results of Haldane for the lattice case in our formulation.

In particular, we used the partition functions to work out the magnetic susceptibility through the thermodynamic potential $\Omega$.

As susceptibility is in general given by

$$\chi = -\frac{\partial^2 \Omega}{\partial H^2} = \frac{\partial^2 \ln Z}{\beta \partial H^2},$$

(4.101)

Using $h = \beta H/2$ and expanding $Z$ to $Z_0 + Z_1 + Z_2 + \cdots$, the susceptibility can be written as

$$\chi = \frac{\beta Z''_0}{4 Z_0} \left( 1 + \left( \frac{Z''_1}{Z''_0} - \frac{Z_1}{Z_0} \right) + \left( \frac{Z''_2}{Z''_0} - \frac{Z_2}{Z_0} - \frac{Z_1}{Z_0} \left( \frac{Z''_1}{Z''_0} - \frac{Z_1}{Z_0} \right) \right) + \cdots \right)$$

$$= \chi_0 (1 + \frac{\chi_1}{\chi_0} + \frac{\chi_2}{\chi_0} + \cdots),$$

(4.102)

in which

$$\chi_0 = \frac{\beta Z''_0}{4 Z_0}, \quad \frac{\chi_1}{\chi_0} = \frac{Z''_1}{Z''_0} - \frac{Z_1}{Z_0}, \quad \frac{\chi_2}{\chi_0} = \frac{Z''_2}{Z''_0} - \frac{Z_2}{Z_0} - \frac{Z_1}{Z_0} \frac{\chi_1}{\chi_0}. \quad (4.103)$$

One can see that if the partition function for the lattice case is the $N$th power of the partition function for the single impurity model, the susceptibility of the lattice model is merely $N$ times the susceptibility of the single impurity case, at least in the perturbation of small $\Delta$ or $V$.

This indicates that the formulation makes no provision for any form of coherent or cooperative phenomenon. Putting it another way, no coherence can be found from investigating the energy levels of the lattice system, in terms of the individual
particles at each site of the lattice. The result indicates that the hybridization be­
tween carriers and localized particles must be quite strong if collective phenomenon
arises in the periodic Anderson model. The perturbative approach to the atomic
limit would therefore become invalid. In the next chapter, we will focus on another
approach which does not impose any condition on $V$ and hence should provide a
better understanding of the model in the large $V$ limit.
CHAPTER 5

Introduction to Canonical Transformation

Canonical transformation is normally considered as a tool by which one can change the basis of operators or vectors from one representation to another. A simple example of representation in our day to day experience is real space in which the actual coordinate of any three dimensional object is the basis vector of the object. Another example of a representation commonly used in physics is momentum space whereby all three dimensional vectors are described by their momentum in three orthonormal directions, instead of their cartesian coordinates. The momentum space representation in many cases yields a great deal of information about a system which would otherwise have been quite difficult to retrieve from the real space representation. This is why both real space and momentum space representations are commonly used in many formulations in physics. The relationship between the two representations can be established by a canonical transformation, with which one can change the basis of a vector from one representation to another. The famous canonical transformation between real and momentum space in this case is the Fourier Transform.

In the past century of physics, canonical transformation often provides a simpler alternative to a complex Hamiltonian which is not obvious to solve. One of the earliest applications of transformation is the meson-nucleon problem by Wentzel [87] in 1943. A later typical example is the Lee, Low Pine transformation [60] for the phonon interaction with charges in a lattice model in 1953. Another example is the displacement operator in quantum mechanics in which the translation and rotation of a particle can be written as the unitary transformation of the corresponding operators. The most famous application of transformation in solid state physics certainly is superconductivity, in which the attractive electron-electron interaction is established from Fröhlich’s proposal [35] by using unitary transformation from
the original electron-phonon interaction.

With the Single Impurity Anderson Model (SIAM) or Periodic Anderson Model (PAM), canonical transformation is commonly used even though the transformation does in fact make the simple Hamiltonian more complex. However, the physics from the outcome of the transformation is more clear and pronounced. In real space for instance the operators of the Hamiltonian are written in terms of site indices and so it is easy to simplify the model, and picture how electrons are moving from one place to another. One can assume, in this case, the conduction electrons are moving slowly from site to site, and only the nearest neighbour hopping and interaction are important. This simplifies the interactions of operators and turns a complicated many body system into a one or two body system. On the contrary, momentum space allows one to assume that the conduction electrons are moving with the same speed and hence the momentum index \( k \) in all free electron operators are approximately the same. This simplifies the kinetic energy term in the Hamiltonian and reduces the number of degrees of freedom of the equation dramatically.

The most relevant transformation of the Anderson model is undoubtedly the canonical transformation by Schrieffer and Wolff \(^1\) who in 1966 discovered from the transformation a spin interaction term which is not explicit from the original Hamiltonian. Using this term, they related the Anderson model to the Kondo model and established equations to transform from one to the other. The physical picture of the transformation with which one could understand the origin of the spin interaction is however not provided. Nowadays, there are many theories to explain the spin interactions in this respect for the Anderson Model, but the physics of the transformation for the Anderson Model is still not well explained.

After various attempts to solve the PAM using perturbation theory in the last two chapters, we will look into the same model from a different angle, using canonical transformation theory and endeavour to explain the physics of the transformation.

## 5.1 General Theory of Transformation

To study the general properties of transformation from one representation to another, we need to write them in a symbolic form, describing the transformation as a result of the action of an operator,

\[
\langle \xi | a \rangle = S(\xi, E_n) \langle E_n | a \rangle
\]

\(^1\)Because of the frequent citation of their single work on this subject, we shall always refer, hereafter, to the result of the transformation by Schrieffer and Wolff to Ref. [78] without explicit citing it.
where we must regard \( S(\xi, E_n) \) to be a matrix with a continuously varying first
index and a discrete second index. The right-hand side of this equation must then
be understood to be the product of the matrix \( S(\xi, E_n) \) with the column matrix
\( \langle E_n | a \rangle \). The transformation in Eq. (5.1) can therefore be written in the form

\[
\langle \xi | b \rangle = S_{\xi,p} \langle p | b \rangle
\]  

The first transformation in Eq. (5.1) is a canonical transformation from one set
of variables \( E_n \) to another set of independent variables \( \xi \). Similarly, Eq. (5.2) is
another canonical transformation from the variable \( p \) to the variable \( \xi \).

We can also define the inverse of Eq. (5.2) as

\[
\langle p | b \rangle = S^{-1}_{\xi,p} \langle \xi | b \rangle
\]  

From the transformation, we have

\[
\langle p | b \rangle = \int d\xi \langle p | \xi \rangle \langle \xi | b \rangle = \int d\xi \langle \xi | p \rangle^\dagger \langle \xi | b \rangle
\]  

Hence,

\[
S^{-1}_{\xi,p} = S^\dagger_{\xi,p}
\]  

or

\[
S^\dagger S = 1
\]  

An operator satisfying condition in Eq. (5.6) is called an *unitary operator* and
the canonical transformations are generated by unitary operators.

In general, we can apply the canonical transformation to any function \( \psi \) by
means of an unitary operator \( S \), written as \( \psi = S\psi \). When we use this equation
to transform the wave functions from one set of variables to another, we must
simultaneously transform all operators. For instance, if an operator \( F_w \) acts upon
a function \( \psi \) in such a way that

\[
\psi' = F_w \psi
\]  

we can transform this equation using the unitary operator \( S \)

\[
S\psi' = SF_w S^{-1} S\psi
\]  

Writing it in terms of \( \phi \),

\[
\phi' = F_\phi \phi,
\]  

where

\[
F_\phi = SF_w S^{-1}
\]
is the operator acting upon $\phi$. Therefore, the transformation of an operator is different to the transformation of a state, but both of them are required in the transformation.

In many cases, the transformation of the Hamiltonian, consisting of a finite combination of operators, is more important than that of the states since the energy content of a system is unaffected by the unitary transformation applied\(^\dagger\). Eq. (5.10) will therefore be of more interest to us, if $H_\psi$ is the Hamiltonian of the system. There are in general two ways of carrying out the transformation in the form of Eq. (5.10). The first is to replace the operators in the Hamiltonian with the transformed version, based on the unitary properties that

\[ S^\dagger(A_1 + A_2)S = S^\dagger A_1 S + S^\dagger A_2 S \quad \text{and} \quad S^\dagger(A_1 A_2)S = S^\dagger A_1 S \cdot S^\dagger B_2 S \]  

(5.11)

In this way, the transformation of any complicated operator in the Hamiltonian, which is represented by a polynomial or power expansion of simpler operator $A_i$, may be written as the transformation of the individual operators $A_i$ in the same functional form of the Hamiltonian.

\[ S^\dagger H(A_i)S = H(S^\dagger AS) \]  

(5.12)

The next step is to expand the unitary transformation $S$ in terms of the set of operators $A_i$ already in the Hamiltonian, so that one can write $S$ as a function of $A_i$. Any transformation of the simple operator $A_i$ will now be written as

\[ S^\dagger(A_i)A_iS(A_i) = f(A_i) \]  

(5.13)

Accordingly, the transformed Hamiltonian $\tilde{H}$ is of a functional form which will be different from that of $H$, but is represented in the same set of variables $A_i$. The Schrieffer and Wolff transformation and the transformation used by Fröhlich in superconductivity [35] for example, belong to this kind of representation.

An alternative is to introduce a new set of dynamic variables $A'_i$, which are defined by the unitary transformation $A'_i = S^\dagger A_i S$ or the inverse $A_i = S^\dagger A'_i S$. The difference between these two representations is that the unitary transformation $S$ in the first case is written as a function of $A_i$, whereas in the second case it is written as a function of $A'_i$. Substituting the second case back into the Hamiltonian, we can easily see that the Hamiltonian is indeed not transformed itself, but its variables $A_i$ are replaced by the new set of variables $A'_i$

\[ H(A_i) = H(S^\dagger(A'_i)A'_iS(A'_i)) \]  

(5.14)

\(^\dagger\)This is mostly the case since the energy of a system is one of the key observables. The wave function is a complex quantity which usually provides a means of finding out other properties of the model.
The functional form of the Hamiltonian with these new variables is normally different to that of the original ones, in order to bring out the characteristics of the system which are usually hidden in the original form.

The two formulations are however identical, only written in two different ways. The first case has the functional form of the Hamiltonian altered to keep the operators $A_i$ unchanged, whereas the second case keeps the functional form of the Hamiltonian relatively the same but written in terms of a new set of transformed operators $A'_i$. The Bogoliubov transformation [19] in the theory of superconductivity for example belongs to this kind of transformation.

The unitary transformation, satisfying Eq. (5.6), can also be written in the form $S = e^{ia}$ where $a$ is a Hermitian operator or an arbitrary real function in the same variables as the wave function. The unitary transformation changes the form, of the phase, of the wave functions, but it does not change the magnitude of the independent variables. This is why unitary transformation is also called phase transformation.

Since physical quantity is real, it can be represented by an infinite number of operators, differing from one another by unitary transformation. In other words, the operators $A$ and $A'$ correspond to the same physical quantity if

$$A' = SAS^{-1}, \quad \text{and} \quad SS^{-1} = 1 \quad (5.15)$$

as the change in phase due to the unitary transformation has no effect to the physical quantity of the variables.

Therefore, the physical quantities of operators remain invariant under unitary transformations. This is only one of the many properties of unitary transformations. In the following, we list the important properties of the unitary transformation.

- It preserves the linearity and hermiticity of operators. If $A$ and $B$ are two operators and $C = A + B$,

$$SCS^{-1} = SAS^{-1} + SBS^{-1}$$

$$\bar{C} = \bar{A} + \bar{B} \quad (5.16)$$

where $\bar{X}$ represents the unitary transformation of any operator $X$. Also

$$(SAS^\dagger)^\dagger = S^\dagger AS \quad \text{if} \quad A = A^\dagger \quad (5.17)$$

- It preserves the scalar product of state vectors which can be proven in the following

$$\langle \tilde{\Psi}_1 | \tilde{\Psi}_2 \rangle = \langle S\Psi_1 | S\Psi_1 \rangle = \langle \Psi_1 | S^\dagger S | \Psi_2 \rangle = \langle \Psi_1 | \Psi_2 \rangle \quad (5.18)$$
where $\tilde{\Psi}$ is the unitary transformation of any wave function $\Psi$. Similarly the matrix elements of operators would not change under unitary transformation. It follows from:

$$\langle \tilde{\Psi}_1 | \hat{A} | \tilde{\Psi}_2 \rangle = \langle S \Psi_1 | S A S^{-1} | S \Psi_1 \rangle = \langle \Psi_1 | S^\dagger S A S^{-1} S | \Psi_2 \rangle = \langle \Psi_1 | A | \Psi_2 \rangle \quad (5.19)$$

- It preserves the spectrum of eigenvalues. If $H \phi_n = E_n \phi_n$, multiply both sides by $S$ and get

$$SH(S^\dagger S)\phi_n = E_n S \phi_n$$
$$\hat{H} \tilde{\phi}_n = E_n \tilde{\phi}_n \quad (5.20)$$

- It preserves commutation and anticommutation relations. It follows from the following. If $C = [A, B]_\pm = AB \pm BA$ and multiply by $S$ and $S^\dagger$ on the left and right respectively, one gets

$$S(AB \pm BA)S^\dagger = SCS^\dagger$$
$$SAS^\dagger SBS^\dagger S = SBS^\dagger SAS^\dagger = \hat{C}$$
$$\hat{A}\hat{B} \pm \hat{B}\hat{A} = \hat{C}$$

using the fact that $SS^\dagger = 1$ for unitary transformation.

- The inverse of $S$ is always equal to its Hermitian conjugate, from the definition of unitary transformation. This makes the process of finding the inverse of $S$ become trivial, as opposed to some other types of transformations which require the help of numerical computation.

So, why would unitary transformation be so popular in theoretical physics and applied science? To answer this question, we will have to look back to the past applications of unitary transformation and find out the advantage it offers to the corresponding applications. In general, one would find that the transformed Hamiltonian $\hat{H}$ is simpler than the original one, in the sense that it would be close to completely diagonal, if not exactly diagonal. The dominant eigenvalues of the Hamiltonian can then be determined from simple algebraic manipulation or approximation. These eigenvalues in general show the ground state and other excited state energies of the system. One example of this is the Bogoliubov transformation in the theory of superconductivity [19] which shows the ground state of the superconductive phase directly out of the eigenvalues of the newly transformed Hamiltonian. More details will be discussed in the next section.
Another reason for an unitary transformation is to get a deeper physical insight into the problem, since the transformed Hamiltonian may reveal the appropriate independent dynamic subsystems. One example of this is the attraction between repulsive electrons in the theory of superconductivity, discovered by Fröhlich [34] initially in 1950 and then reformulated by himself using canonical transformation in 1952 [35]. Another example is the particle dressing in a phononic system [60] in which a transformation would change the "electronic coordinate" of the original Hamiltonian to the coordinate of the newly "dressed" particle.

Finally, it should be emphasized that unitary transformations as such should not be mistaken as an approximation method. It is always an exact procedure although it may be supplemented by approximated methods after being applied.

5.2 Renowned Applications to Condensed Matter Systems

In the rest of this chapter, we will demonstrate the effectiveness of canonical transformation to the understanding of a system, by introducing three examples which have had significant impact on the course of science. Even though they are in different areas of condensed matter physics, the canonical transformations in all three cases bring out the characteristics of the systems which are not at all obvious from the original form of the Hamiltonian. The first example is in the theory of superconductivity in which two unitary transformations have played a vital role. The first conveys the original electron-phonon interaction as an effective electron-electron coupling which uncovered the attractive interaction between electrons, as first explained by Fröhlich [35]. The Hamiltonian can also be diagonalized using what is known as the Bogoliubov transformation [19, 83] after incorporating the attraction between electrons in it. The second example is in the high temperature superconductivity, regarding the derivation of the $t - J$ model from a one band Hubbard model. The transformation brings out the magnetic coupling between the electrons within the same energy band of the system. The third example is in the quantum theory of magnetism, in which Schrieffer and Wolff used a canonical transformation to prove the existence of spin interaction in the single impurity Anderson model. This established the relationship between the Anderson model and the Kondo model.
5.2. Superconductivity

The theory of electric conductivity of metals was first developed by Bloch [18], in which the resistivity is caused by the scattering of electrons from the lattice vibrations. This scattering involves the creation of phonons when an electron transfers part of its momentum and energy to an ion in the lattice during scattering. This phonon with a certain energy and momentum then travels along the lattice until it hits another electron to which all or some of its energy and momentum are transferred. This emission and absorption of quanta by electrons connects the electrons and hence the self-energy of the electrons will no longer be the same as the bare electrons in vacuum. Putting it another way, an electron produces a lattice deformation which in turn reacts on the electron itself, and therefore the self-energy will not be the same as if it was in a vacuum. Fröhlich [34] was the first to report that this kind of interaction between electrons through the lattice vibration can be attractive given certain limited conditions. Bardeen and Pine [13] provided a more complete analysis, following the idea of Fröhlich, which contains the Coulomb effects in the interactions between electrons. All of these attempts were however based essentially on the self-energy of the electrons. It was later recognized that a true interaction between electrons indeed played a more essential role. This attraction is nevertheless the corner stone in the theory of superconductivity. In Fröhlich's second paper [35], he used a canonical transformation to prove his point which we will discuss in the following. Our approach is more direct and simple, following that of Wagner [85].

Fröhlich started with a hamiltonian in a second quantized form with respect to the electronic system,

$$H = \sum_k \hbar \omega_k c_k^\dagger c_k + \sum_q \hbar \Omega_q (b_q^\dagger b_q + \frac{1}{2}) + \sum_{k,q} [g_q^* b_q^\dagger c_k^\dagger c_{k+q} + \text{H.C.}], \quad (5.21)$$

in which $c_k^\dagger (c_k)$ is the electron creation (destruction) operator, $b_k^\dagger (b_k)$ is the phonon creation (destruction) operator. $\omega_k$ is the dispersion of conduction electrons while $\Omega_q$ is the dispersion of the phonon field. $g_q$ is the coupling parameter between electrons and phonons which satisfies $g_q^* = -g_q$, and can be taken as purely imaginary. Spin indices have been neglected for simplicity as the impact of spin to the interaction is minimal.

Using the canonical transformation

$$U = e^S, \quad \text{where} \quad S = -\sum_{k,q} \alpha_{k,q} (g_q^* b_q^\dagger c_k^\dagger c_{k+q} - \text{H.C.}), \quad (5.22)$$

on the Hamiltonian in which $\alpha_{k,q}$ is a variable to be determined later, for removing
certain type of interaction in the transformed Hamiltonian. We get

\[ \tilde{H} = U^\dagger H U \]

\[ = \hbar \sum_k \omega_k c_k^\dagger c_k + \hbar \sum_q \Omega_q (b_q^\dagger b_q + \frac{1}{2}) - \sum_{k,q} W_{ep}(k,q) (g_q^* b_q^\dagger c_{k+q} + \text{H.C.}) \]

\[ + \sum_{k,q} W_{ren}(k,q) |g_q|^2 (c_{k+q}^\dagger c_{k+q} - c_k^\dagger c_k) (b_q^* b_q + \frac{1}{2}) \]

\[ + \sum_{k,q,k'} W_{ee}(k,q) |g_q|^2 c_{k+q}^\dagger c_k^\dagger c_{k'-q}^\dagger c_{k'} + O(g^3), \]

where

\[ W_{ep}(k,q) = -1 + \hbar \alpha_{k,q} (\omega_k - \omega_{k+q} + \Omega_q), \]

\[ W_{ren}(k,q) = \alpha_{k,q} [-2 + \hbar \alpha_{k,q} (\omega_k - \omega_{k+q} + \Omega_q)], \]

\[ W_{ee}(k,q) = \frac{1}{2} \alpha_{k,q} [-2 + \hbar \alpha_{k,q} (\omega_k - \omega_{k+q} + \Omega_q)] \]

\[ + \frac{1}{2} \alpha_{k+q,-q} [-2 + \hbar \alpha_{k+q,-q} (\omega_{k+q} - \omega_k + \Omega_q)], \]

The second and higher order phonon terms are discarded as we are only interested in the first order result. Using the interpretation of Haug [42], the first three terms represent the electrons, phonons and the residual electron-phonon interaction, respectively. The fourth term produces a renormalization of the lattice frequencies and the fifth term is the new electron-electron interaction. Fröhlich chose the \( \alpha_{k,q} \) in \( S \) of Eq. (5.22) to eliminate the electron-phonon interaction term \( W_{ep} \).

\[ \left[ \sum_k \hbar \omega_k c_k^\dagger c_k + \sum_q \hbar \Omega_q (b_q^\dagger b_q + \frac{1}{2}), S \right] = - \sum_{k,q} [g_q^* b_q^\dagger c_{k+q} + \text{H.C.}], \]

This is equivalent to setting \( W_{eq}(k,q) \) to zero which yields

\[ \alpha_{k,q} = \frac{-1}{\hbar (\omega_{k+q} - \omega_k - \Omega_q)}, \]

\[ W_r = \frac{1}{\hbar (\omega_{k+q} - \omega_k - \Omega_q)}, \]

\[ W_{ee} = \frac{\Omega_q}{\hbar (\omega_{k+q} - \omega_k)^2 - \Omega_q^2}, \]

The fact that the coefficient \( W_{ee} \) of the two-electron interaction term can be negative if \( (\omega_{k+q} - \omega_k)^2 < \Omega_q^2 \), implies that an attractive interaction is possible between two electrons. Only the electron pairs by themselves are still inadequate to explain superconductivity since the pairs could have encountered resistance travelling through the lattice of positive ions. However, the two electrons in this interaction are close to the
Fermi surface, for the inequality to be satisfied, and the energy difference between the two electrons has to be smaller than a quantum packet of the phonon energy. This implies that the electrons in this case are incapable of transferring energy to the lattice, and therefore the electrons must carry all of their energy and travel free of scattering in the lattice. In terms of band theory, there are gaps between different energy levels of phonons and within these gaps no phonons can be created or destroyed since phonons cannot exist with a fraction of a quantum packet. The electron pairs formed due to the attractive interaction mentioned above happen to fall within this gap, to satisfy the inequality equation. It is therefore energetically favourable for electrons to travel in pairs without encountering any resistance.

From the calculation, electrons form pairs in the momentum space, not in real space. This means two electrons can couple even if they are physically quite far apart. One might wonder how this is possible. In terms of the Hamiltonian, the energy gain from the negative potential of an electron pair comes from the coherence of the electron wave functions with lattice vibrations of short wavelength. The electron pair is synchronous in their wave function when they pass by each other and it is this synchronization that causes the attractive interaction. The bound state formed in this way would not represent a condensation in real space and the physical properties are therefore very different to that of a conventional condensate.

Even though the theory of Fröhlich and Bardeen explained the formation of pairs of electrons, it was not explicitly stated that the pair must also have opposite momentum, until Bardeen, Cooper and Schrieffer [14] who specifically assumed that the total momentum must be small so that the individual momentum of each electron in the pair will satisfy the inequality. When the total momentum is zero, it maximizes the transitions between electrons and ensures that if one electron of a pair is in the pair state, so is the electron with an opposite momentum. Cooper also proved in his landmark paper [28] that for a negative interaction between two electrons, a bound state must form no matter how weak the interaction is. All of these laid the path to the final theory of superconductivity.

Fröhlich theory [34] was a great break-through which successfully explained the isotope effect, where the superconductivity transition temperature \( T_c \) of different isotopes of the same element varies with the ionic mass \( M \) as \( T_c \propto M^{-1/2} \). This result indicates that the dynamics of the ionic cores affects the superconducting state, even though the ions are not especially important in the normal state. It also suggests that electron-phonon interactions are primarily responsible for superconductivity.

However, all theories based upon the self-energy of electrons or the electron-
phonon interactions alone do not yield the correct energy difference between the normal and the superconducting phases. The calculated energy is far too large compared with the observed value. This puzzle was later resolved by Bardeen, Cooper and Schrieffer [15] when they considered the lower bound state in superconductivity is mainly due to the attractive interaction between electron pairs, rather than the self-energy of electrons or direct phonon interaction.

The above discussion essentially concludes that two particles of opposite momenta near the Fermi surface will form a bound pair in momentum space, as long as the interaction between the electrons is attractive. This, as a result, would lower the energy of the system, leading the original unperturbed ground state to become unstable. Unfortunately, Cooper's model is restricted to only two particles and therefore incapable of deriving the new ground state with many body interaction. To this end, we will have to turn to the Bogoliubov canonical transformation which provides the many-body effect of the interacting Hamiltonian using a quasi-particle picture.

The Bogoliubov canonical transformation is versatile. When it is applied to an interacting Bose system on one hand, it explains the phenomenon of superfluidity in terms of quasi-particles if the interparticle interaction is repulsive. This repulsive interaction is quite reasonable between bosons since an attractive interaction would lead to a totally collapsed ground state which is completely different to the non-interacting ground state. On the other hand, when it is applied to an interacting Fermi system such as electrons in metal, it produces the result of superconductivity. We will follow the approach of Beliaev [17] in the following to first investigate the system of free Fermions, and then introduce interaction between the Fermions.

The Hamiltonian of free non-interacting Fermi particles can be written as

$$H_0 = \sum_k \epsilon_k a_k^\dagger a_k,$$

in which $a_k^\dagger$ and $a_k$ are fermionic creation and annihilation operators following the usual anti-commutation rule. The energy of the particles $\epsilon_k$ is symmetric with momentum $k$. The total number of particles in the system is fixed, as no particle exchange is allowed in the Hamiltonian. To allow for an open system environment similar to the one in real metals, it is more realistic to consider a grand canonical ensemble as explained in chapter 3 in which the Hamiltonian should be modified to

$$H_0 = \sum_k \epsilon_k a_k^\dagger a_k - \mu N,$$

where $N$ is the total number operator $N = \sum_k a_k^\dagger a_k$ and $\mu$ is the chemical potential. The chemical potential can indeed be considered as a Lagrange multiplier to be
5.2. Renowned Applications to Condensed Matter Systems

determined by the requirement that

$$\langle N \rangle = N_0 = \text{average number of particles},$$  \hspace{1cm} (5.34)

This requirement should fix $\mu$ which can then be used in the calculation of the thermodynamic properties.

In the ground state when all particle states are filled up to the Fermi energy level, or the chemical potential in the non-interacting system, we can consider a new quasi-particle operator $\alpha^\dagger_k$ which creates a particle outside the Fermi surface, or a hole inside the Fermi sphere.

$$\alpha^\dagger_k = \begin{cases} a_k^\dagger & \epsilon_k > \mu, \\ a_k & \epsilon_k < \mu. \end{cases}$$  \hspace{1cm} (5.35)

This quasi-particle operator, like the original operator $a_k^\dagger$, satisfies anti-commutation rules. The momentum index $k$ of this new quasi-particle corresponds to the momentum of the original particle when $\epsilon_k > \mu$, or the momentum of a hole when $\epsilon_k < \mu$. In terms of this representation, the Hamiltonian becomes

$$H_0 = \sum_{k,\epsilon_k<\mu}(\epsilon_k - \mu)(1 - \alpha^\dagger_{-k}\alpha_{-k}) + \sum_{k,\epsilon_k>\mu}\alpha^\dagger_k\alpha_k$$

$$= E_F + \sum_{k}E_k\alpha^\dagger_k\alpha_k$$  \hspace{1cm} (5.36)

where $E_F = \sum_{k,\epsilon_k<\mu}(\epsilon_k - \mu)$ is the total ground state energy and $E_k = |\epsilon_k - \mu|$ is the quasi-particle energy. This shows quite clearly that each quasi-particle changes the ground state energy content of the system $\sum_{k,\epsilon_k<\mu}(\epsilon_k - \mu)$ by an amount $E_k$. To be more precise, we should write $|\text{vac}\rangle$ as the true vacuum state of system without any particles and $|0\rangle$ as the “vacuum” state of quasi-particles, so that

$$a_k|\text{vac}\rangle = 0, \quad \alpha_k|0\rangle = 0.$$  \hspace{1cm} (5.37)

This shows the resemblance between the particle operator $a_k$ to the vacuum state and the quasi-particle operator $\alpha_k$ to the ground state. The corresponding Hamiltonian is therefore

$$H_0|0\rangle = E_F|0\rangle, \quad \text{and} \quad H_0\alpha^\dagger_k|0\rangle = (E_F + E_k)\alpha^\dagger_k|0\rangle$$  \hspace{1cm} (5.38)

in which the ground state of the quasi-particles is used, instead of the true vacuum state. $E_k$ is therefore the energy of the quasi-particle of momentum $k$. The total number operator $N$ can also be written in terms of this new quasi-particle picture and the average of $N$ with respect to $|0\rangle$, $(0|N|0)$, should give the total number of
particles in the non-interacting ground state. In this case, the chemical potential is identical to the Fermi energy and the number of particles \( N \) should be equal to the number of available states below the Fermi surface. One might say that there is no real advantage of using quasi-particles in this simple case, as the original fermionic operator is clear enough to provide the overall picture of the system. The advantage however comes from the interacting case which we will consider now.

In a system of a large number of weakly interacting fermions, the Hamiltonian of a grand canonical ensemble can be written as:

\[
H = \sum_{k, \sigma} (\epsilon_k - \mu) a_{k, \sigma}^\dagger a_{k, \sigma} - \frac{1}{2} \sum_{\text{all} k} \langle k_1, k_2 | G | k_1', k_2' \rangle a_{k_1, \sigma_1}^\dagger a_{k_2, \sigma_2}^\dagger a_{k_2', \sigma_2'} a_{k_1', \sigma_1'} \tag{5.39}
\]

where \( G \) is the two particle interaction operator, or the scattering matrix. The four momentum indices in the summation are not entirely independent, due to the conservation of momentum, which must satisfy \( k_1 + k_2 = k_1' + k_2' \). The minus sign in front of this interaction term is to reflect what Fröhlich and Bardeen considered the attractive interaction between a pair of electrons. The findings of Fröhlich and Cooper suggested that there exists a strong correlation between particles of opposite momenta and as a result they form a bound state in the momentum state. Using this as a hint, one would therefore consider a canonical transformation that could turn a homogeneous system into a system consisting of two different kinds of quasi-particles. Quasi-particles of these two kinds have a tendency to form a pair in the momentum space to lower the overall energy of the system.

Similar to the approach of the non-interacting system, we shall call these two kinds of quasi-particles \( \alpha_k \) and \( \beta_k \) which are given by

\[
\begin{align*}
\alpha_k &= u_k a_{k, 1/2} - v_k a_{-k, -1/2}^\dagger, \\
\beta_k &= u_k a_{-k, -1/2} + v_k a_{k, 1/2}^\dagger.
\end{align*}
\tag{5.40}
\]

with the relationship between the coefficients

\[
u_k^2 + v_k^2 = 1, \tag{5.42}
\]
satisfied. These two coefficients are real for all \( k \) and symmetric with respect to \( k \), \( i.e. \ u_k = u_{-k} \) and \( v_k = -v_{-k} \). The new quasi-particles satisfy anti-commutation rules among themselves and between the two different kinds.

\[
\{ \alpha_k, \alpha_{k'} \} = \{ \beta_k, \beta_{k'} \} = \{ \alpha_k, \beta_{k'} \} = \{ \alpha_{k'}, \beta_k \}. \tag{5.43}
\]

But

\[
\{ \alpha_k, \alpha_{k'}^\dagger \} = \{ \beta_k, \beta_{k'}^\dagger \} = \delta_{k,k'}, \tag{5.44}
\]
If we choose for example

\[
    u_k = \begin{cases} 
    1 & \epsilon_k > \mu, \\
    0 & \epsilon_k < \mu 
    \end{cases} \\
    v_k = \begin{cases} 
    0 & \epsilon_k > \mu, \\
    1 & \epsilon_k < \mu 
    \end{cases} \quad (5.45)
\]

this transformation will be essentially the same as the one defined in the non-interacting case. It therefore suggests that \( \alpha_k \) and \( \beta_k \) in the interacting model are two quasi-particles of a similar nature to the quasi-particle in the non-interacting case. To determine the values of \( u_k \) and \( v_k \), we first need to find the Hamiltonian in terms of these new quasi-particles.

\[
    H = U + H_{11} + H_{20} + H_{\text{int}} \quad (5.46)
\]

in which

\[
    U = \sum_k 2(\epsilon_k - \mu - \frac{1}{2} \sum_{k'} \langle k, k' | \bar{G} | k', k \rangle v_{k'}^2) u_k^2 - \sum_{k,k'} \langle k, -k|G|k', k'\rangle u_{k'} v_k u_k v_k, \quad (5.47)
\]

\[
    H_{11} = \sum_k \left( (\epsilon_k - \mu - \sum_{k'} \langle k, k' | \bar{G} | k', k \rangle v_{k'}^2)(u_k^2 - v_k^2) + 2 \sum_{k'} \langle k, -k|G|-k', k'\rangle u_{k'} v_k u_k v_k \right) (\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k), \quad (5.48)
\]

\[
    H_{20} = \sum_k \left( 2(\epsilon_k - \mu - \sum_{k'} \langle k, k' | \bar{G} | k', k \rangle v_{k'}^2) u_k v_k \\
    - \sum_{k'} \langle k, -k|G|k', k'\rangle u_{k'} v_{k'} (u_k^2 - v_k^2) \right) (\alpha_k \beta_k^\dagger + \beta_k \alpha_k), \quad (5.49)
\]

and \( \langle k, k' | \bar{G} | k', k \rangle \) is defined as

\[
    \langle k, -k' | G | -k', k \rangle + \langle k, k' | G | k', k \rangle - \langle k, k' | G | k, k' \rangle, \quad (5.50)
\]

\( H_{\text{int}} \) is a collection of various products of four particle operators. As discussed in [17], one can ignore the term \( H_{\text{int}} \) without altering the physics of the model significantly. To diagonalize this new Hamiltonian, we simply set the coefficient of \( H_{20} \) to zero which yields

\[
    2\xi_k u_k v_k - \Delta_k (u_k^2 - v_k^2) = 0, \quad (5.51)
\]

in which we introduced the measure of energy relative to the chemical potential \( \xi_k \) and the energy gap \( \Delta_k \) between the quasi-particle ground state and the non-interacting ground state. The concept of single quasi-particle energy \( \tilde{\epsilon}_k \) is needed
in the formulation of chemical potential \( \xi_k \), as shown below.

\[
\bar{\epsilon}_k = \epsilon_k - \sum_{k'} \langle k, k' | \hat{G} | k', k \rangle u_{k'}^2,
\]

\[
\xi_k = \bar{\epsilon}_k - \mu,
\]

\[
\Delta_k = \sum_{k'} \langle k, -k | G | -k', k' \rangle u_{k'} v_{k'}.
\]

Together with the normalization Eq. (5.42) condition, we can write

\[
u_k = \cos \chi, \quad v_k = \sin \chi
\]

so that Eq. (5.51) is equivalent to

\[
\tan 2\chi = \frac{\Delta_k}{\xi_k}
\]

In order to solve for \( \Delta_k \) from these equations, we first make an approximation which was first observed from Cooper [28] as mentioned before that, a pair of electrons will form a bound state regardless of how small the attraction is. We can therefore assume \( G \) as

\[
G = \begin{cases} 
G_0 & \text{when } \mu - \omega < \bar{\epsilon}_k, \\
0 & \text{when } \mu + \omega > \bar{\epsilon}_k.
\end{cases}
\]

where \( \omega \) is the cut-off parameter in energy for \( G \) to be non-zero. In fact, \( \omega \) in real physical processes corresponds to the average frequency of the phonons in the lattice below which the amount of phonon excitation is minimal. Besides, the constant \( G_0 \) is \( k \) independent which implies that the set of equations we want to solve can be further simplified. Using this in the definition of \( \Delta_k \) function, together with the substitution of \( u_k \) and \( v_k \), we have for \( \Delta_k = \Delta \) inside the cut-off range,

\[
\Delta = \frac{1}{2} \sum_{k'} G \Delta \frac{G}{\sqrt{\xi_k^2 + \Delta^2}}
\]

in which the summation \( \sum'_{k'} \) is restricted over the non-zero region of \( G \). This equation has two solutions, one being a trivial solution \( \Delta = 0 \) which corresponds to \( u_k v_k = 0 \). The solutions of \( u_k \) and \( v_k \) in this case are therefore given by Eq. (5.45). The definition of quasi-particles to which this solution leads is identical to that for the non-interacting case considered before. It is normally called the normal solution.

For the non-trivial solution, we will first need to turn the summation into an integration

\[
1 = \frac{1}{2} \sum'_{k'} \frac{G}{\sqrt{\xi_k^2 + \Delta^2}} = \frac{G}{2} \int_{\mu - \omega}^{\mu + \omega} \frac{\rho(\bar{\epsilon}) \bar{\epsilon} \, d\bar{\epsilon}}{\sqrt{(\bar{\epsilon} - \mu)^2 + \Delta^2}}
\]
where we have used the density of states $\rho$ in terms of energy to convert the summation from momentum $k$ to energy $\epsilon$. Next, we suppose that $\Delta$ is very small. This can be verified later, that significant contributions to the integral come from the Fermi surface. This implies $\epsilon \approx \mu$ and so $\rho(\epsilon) \approx \rho$ is a constant over the limit of the integration. The non-trivial solution then becomes

$$1 = \frac{G\rho}{2} \int_{\mu-\omega}^{\mu+\omega} \frac{d\epsilon}{\sqrt{(\epsilon - \mu)^2 + \Delta^2}} = G\rho \sinh^{-1} \frac{\omega}{\Delta}$$  \hspace{1cm} (5.60)

If we write $\eta = 1/G\rho$, the solution for $\Delta$ becomes

$$\Delta = \frac{\omega}{\sinh \eta}$$ \hspace{1cm} (5.61)

In the case of weakly interacting system, $G\rho \ll 1$ and so $\eta \gg 1$ which allows us to write

$$\Delta = \omega e^{-\eta}$$ \hspace{1cm} (5.62)

This is what is normally called, \textit{superconducting} solution. The transformed Hamiltonian now takes the form

$$H = U + \sum_k E_k(\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k),$$ \hspace{1cm} (5.63)

and the energy of the quasi-particles is

$$E_k = \xi_k(u_k^2 - v_k^2) + 2\Delta_k u_kv_k = \sqrt{\xi_k^2 + \Delta_k^2},$$ \hspace{1cm} (5.64)

From these two equations together with the equations for a superconducting state, one can derive all the equations necessary for the BCS theory [15]. Before we leave for the next example, we would like to make some remarks about this transformation.

First of all, the number of particles in the ground state can be derived as follows.

$$\langle N \rangle = \langle \sum_{k,\sigma} a^\dagger_{k,\sigma} a_{k,\sigma} \rangle = 2 \sum_k u_k^2 = 2 \sum_k \sin^2 \chi,$$ \hspace{1cm} (5.65)

which shows a round smeared transition at the Fermi surface, instead of a sudden jump in the normal solution of Eq. (5.45) as shown in figure 5.1. The width of the transition is measured by the parameter $\Delta$. Bearing in mind that this is still at $T = 0$, it indicates that some of the electrons (above the Fermi surface) and holes (below the Fermi surface) manage to escape from the Fermi surface and settle at a slightly higher superconducting state.
Secondly, if one calculates the energy difference between the superconducting state and the normal state, one would get [17]

\[ E_{\text{super}} = E_{\text{normal}} - 2\rho\omega^2 e^{-2\eta} \]  

(5.66)

which implies the ground state energy for the superconducting solution is lower than that of the normal solution.

Thirdly, the condition we chose in the canonical transformation to eliminate the cross terms in fact also minimizes the vacuum state energy or zero-point energy \( U \) in Eq. (5.46) which can be verified by substituting the solution of \( \tan 2\chi \) from Eq. (5.56) to the partial derivative of \( U \) with respect to \( \chi_k, \partial U/\partial \chi_k \). The result is zero which can be confirmed as a minimum point.

Finally, the excitation spectrum of the quasi-particle from the transformation indicates that there is a minimum close to \( k' = k'_p \) with the value of \( \Delta \) above zero. The quasi-particle therefore has a "gap" of amount \( \Delta \) between the non-interacting ground state and the lowest excited level. As a result, it is necessary to have at least an excitation energy of \( \Delta \) to create a quasi-particle which explains the phenomenon of superconductivity. Also, the excited states of the quasi-particle originates from

\[ \begin{align*}
  &\text{Figure 5.1: The average number of particles in the superconducting state (solid line), as opposite to that in the normal state (dashed line).}
\end{align*} \]
5.2. Renowned Applications to Condensed Matter Systems

breaking up one or more pairs of electrons in the superconducting ground state, and so the excitation energy is in fact \(2\Delta\).

5.2.2 The \(t - J\) Model

The second example we will discuss below is the famous canonical transformation of the Hubbard Model which shows explicitly the spin interaction between electrons. The result of the transformation is known as the \(t - J\) model, as it contains two important parameters \(t\) and \(J\), instead of \(t\) and \(U\) in the Hubbard Model. \(t\) is the hopping matrix of conduction electrons and relates to the kinetic energy of the model, whereas \(J\) corresponds to the spin interactions among the electrons in the model. The formulation below is mostly based upon the original paper by Chao, Spalek and Oleš [22] who introduced a canonical transformation to remove electron hopping across different energy subbands of the system. As a result, the effective Hamiltonian obtained allows only for hopping within the same band. The longitudinal spin-spin coupling of electrons gives rise to virtual hopping, while the transverse spin-spin coupling yields spin-wave excitation. In fact, the effective Hamiltonian resembles a Heisenberg antiferromagnet in terms of spin-interaction, but the itinerant property of electrons within a given sub-band destroys the anti-ferromagnetic ordering of the electrons as \(t/U\) increases. This is in contrast to the Canonical Perturbation Theory (CPT) we mentioned in chapter 4 where the intra-subband hopping is projected out and only the inter-subband hopping was discussed.

We start with a two-site Hubbard model and later extend to a many sites model.

\[
H = t \sum_{ij,\sigma} c_{i,\sigma}^\dagger c_{j,\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow},
\]

(5.67)

where \(i\) and \(j\) can be any one of the two sites, but never the same, \textit{i.e.} \(i \neq j\). \(c_{i,\sigma}^\dagger\) creates an electron at site \(i\) of spin \(\sigma\) and \(c_{i,\sigma}\) destroys the electron. \(n_{i,\sigma}\) is the number operator of the \(c\)-electrons. As mentioned before, \(t\) is the hopping matrix (in this case a constant) which determines the kinetic energy of the electrons. \(U\) represents the strong Coulomb repulsion between electrons at the same site. This model has been extensively studied in the 1970s and 80s, especially in relation to the subject of high temperature superconductivity. The understanding of it in the two and three dimensional lattice is however still very limited.

The model looks very simple at first glimpse, but the mathematics behind it is very complex. The difficulty in solving this model is mainly due to the on-site interaction in the form of a Dirac delta function between two particles in the second term of the Hamiltonian. This is part of the reason why an exact solution has not
yet been found for the model in two and three dimensions. In order to visualize the
physics behind the Hamiltonian, we need to turn the Hamiltonian into a form that
is more explicit with the electrons hopping within subbands and between subbands.
One way to do so is to introduce projection operators. The projection operator for
each site $i$ of the lattice is defined as

$$\Pi_\sigma (1 - n_{i,\sigma}) + n_{i,\sigma} x = \sum_{m=0}^{2} p(im)x^m, \quad (5.68)$$

in which $p(im)$ projects out the atomic configurations each of which contains $m$
electrons. $p(im)$ is complete and orthogonal, i.e. $\sum_{m=0}^{2} p(im) = 1$ and $p(im)p(in) =$
$p(im)\delta_{mn}$. This allows us to define the projection operators for our two site model,
in terms of $p(im)$,

$$P_1 = p(11)p(21) + p(11)p(20) + p(21)p(10) + p(10)p(20),$$

$$P_2 = p(12)p(21) + p(22)p(11) + p(12)p(20) + p(22)p(10) + p(12)p(22),$$

Similar to $p(im)$, $P_\mu$ are also complete and orthogonal where $\mu = 1, 2$, i.e. $\sum_{\mu} P_\mu = 1$
and $P_\mu P_\nu = P_\mu \delta_{\mu,\nu}$. Applying $P_\mu$ to both sides of the Hamiltonian in Eq. (5.67)

$$P_1 H P_1 = t \sum_{ij,\sigma} (1 - n_{i-\sigma}) c_{i,\sigma}^\dagger c_{j,\sigma} (1 - n_{j,-\sigma}), \quad (5.69)$$

$$P_1 H P_2 = t \sum_{ij,\sigma} (1 - n_{i,-\sigma}) c_{i,\sigma}^\dagger c_{j,\sigma} n_{j,-\sigma}, \quad (5.70)$$

$$P_2 H P_1 = t \sum_{ij,\sigma} n_{i,-\sigma} c_{i,\sigma}^\dagger c_{j,\sigma} (1 - n_{j,-\sigma}), \quad (5.71)$$

$$P_2 H P_2 = t \sum_{ij,\sigma} n_{i,-\sigma} c_{i,\sigma}^\dagger c_{j,\sigma} n_{j,-\sigma} + U \sum_{i} n_{i,\uparrow} n_{i,\downarrow}, \quad (5.72)$$

we obtain a Hamiltonian projected onto four different sub-spaces. The first and the
last of these four equations represent two energy subbands separated roughly by
$U$ when $U \gg t$, whereas the rest of the four equations describe hopping between
these two energy subbands. In other words, we have separated the inter-subband
hopping from the intra-subband hopping of the model. As our objective is to dis-
cover hopping within the same energy subband, the next step is to find a canonical
transformation that will project out this cross subbands hopping. Before starting
the calculation, we rewrite the Hamiltonian as

$$H(\epsilon) = H_0 + \epsilon H_1, \quad (5.73)$$
5.2. Renowned Applications to Condensed Matter Systems

in which

\[ H_0 = P_1 H P_1 + P_2 H P_2, \quad (5.74) \]
\[ H_1 = P_1 H P_2 + P_2 H P_1, \quad (5.75) \]

\( \epsilon \) is formally treated as small except at the end of the calculation when \( \epsilon \) is set to one. The canonical transformation in terms of \( \epsilon \) is

\[ \tilde{H}(\epsilon) = e^{-i\epsilon S} H(\epsilon) e^{i\epsilon S}, \quad (5.76) \]

where \( S \) is determined so that the final Hamiltonian contains no term linear in \( \epsilon \). Expanding the Hamiltonian in order of \( \epsilon \) using the expansion

\[ e^{i\epsilon S} = 1 + i\epsilon S - \frac{1}{2} \epsilon^2 S^2 + \cdots, \quad (5.77) \]

we obtain \( \tilde{H} \) to the second order of \( \epsilon \) as

\[ \tilde{H}(\epsilon) = H_0 + \epsilon (H_1 + i[H_0, S]) + \frac{1}{2} \epsilon^2 (i[H_1, S] - [H_0, S]) \]

(5.78)

The term linear in \( \epsilon \) vanishes if

\[ H_1 = i[H_0, S] = 0, \quad (5.79) \]

and the Hamiltonian after setting \( \epsilon \) to one becomes

\[ \tilde{H}(\epsilon = 1) = H_0 + \frac{1}{2} i[H_1, S], \quad (5.80) \]

The key to this transformation now falls on the determination of \( S \) from Eq. (5.79). First of all, let us project this condition in Eq. (5.79) onto different subspaces by applying the operator \( P_\mu \) on both sides of the equation.

\[ P_\mu H P_\nu (1 - \delta_{\mu,\nu}) + i P_\mu H P_\mu (P_\mu S P_\nu) - i (P_\mu S P_\nu) P_\nu H P_\nu = 0, \quad (5.81) \]

When \( \mu = \nu \), it reduces to

\[ P_\mu H P_\nu (P_\mu S P_\mu) = (P_\mu S P_\mu) P_\mu H P_\mu, \quad (5.82) \]

which has the general solution of \( P_\mu S P_\mu = \gamma P_\mu \) where \( \gamma \) is an arbitrary constant.

For \( \mu \neq \nu \), Eq. (5.81) becomes

\[ P_\mu H P_\mu (P_\mu S P_\nu) - (P_\mu S P_\nu) P_\mu H P_\nu = i P_\mu H P_\nu, \quad (5.83) \]

If we approximate the operators \( P_\mu H P_\mu \) and \( P_\nu H P_\nu \) by their expectation values, the right hand side of this equation consists of all possible single-hopping processes
between the two energy subbands, multiplied by the mean energy difference between the final and the initial states of the cross subbands hopping. When \( U \gg t \), this energy difference can be well approximated by the separation of the subbands \( U \) and hence,

\[
P_1SP_2 = -\frac{i}{U}P_1HP_2, \quad P_2SP_1 = \frac{i}{U}P_2HP_1,
\]

(5.84)

The effective Hamiltonian \( \hat{H} \) is now given by

\[
\hat{H} = P_1HP_1 + P_2HP_2 - \frac{1}{U}(P_1HP_2HP_1 - P_2HP_1HP_2),
\]

(5.85)

Expanding the projection operators in terms of electron operators using Eqs. (5.69-5.72), we have

\[
\hat{H} = -t \sum'_{ij,\sigma} (1 - n_{i,-\sigma})c_{i,\sigma}^\dagger c_{j,\sigma}(1 - n_{j,-\sigma}) + t \sum'_{ij,\sigma} n_{i,-\sigma}c_{i,\sigma}^\dagger c_{j,\sigma}n_{j,-\sigma}
\]

\[
+ U \sum_i n_{i,+}n_{i,-} - \frac{t^2}{U} \left( (\sum'_{ij,\sigma} (1 - n_{i,-\sigma})c_{i,\sigma}^\dagger c_{j,\sigma}n_{j,-\sigma}) \cdot (\sum'_{ij,\sigma} n_{i,-\sigma}c_{i,\sigma}^\dagger c_{j,\sigma}(1 - n_{j,-\sigma})) - (\sum'_{ij,\sigma} n_{i,-\sigma}c_{i,\sigma}^\dagger c_{j,\sigma}(1 - n_{j,-\sigma})) \cdot (\sum'_{ij,\sigma} (1 - n_{i,-\sigma})c_{i,\sigma}^\dagger c_{j,\sigma}n_{j,-\sigma}) \right),
\]

(5.86)

If we now sum over the whole lattice instead of just over two sites and consider the interaction among the electrons to be the strongest when they are the nearest, the summation over \( i \) and \( j \) can be replaced by a summation over all nearest neighbour sites \( (i, j) \). Also, the result can be simplified by using the dot product of two spin operators

\[
S_i \cdot S_j = S^x_i S^x_j + S^y_i S^y_j + S^z_i S^z_j
\]

\[= \frac{1}{2}(S^+_i S^-_j + S^-_i S^+_j) + S^z_i S^z_j
\]

\[= \frac{1}{2}(c_{i,\uparrow}^\dagger c_{j,\downarrow} c_{j,\uparrow}^\dagger c_{i,\downarrow} + c_{i,\downarrow}^\dagger c_{j,\uparrow} c_{j,\downarrow}^\dagger c_{i,\uparrow}) + \frac{1}{4}(n_{i,\uparrow} + n_{i,\downarrow})(n_{j,\downarrow} + n_{j,\uparrow}),
\]

(5.87)

which turns the effective two site Hamiltonian to

\[
\tilde{H} = t \sum'_{\langle ij \rangle,\sigma} (1 - n_{i,-\sigma})c_{i,\sigma}^\dagger c_{j,\sigma}(1 - n_{j,-\sigma}) + t \sum'_{\langle ij \rangle,\sigma} n_{i,-\sigma}c_{i,\sigma}^\dagger c_{j,\sigma}n_{j,-\sigma}
\]

\[+ U \sum_i n_{i,+}n_{i,-} - \frac{t^2}{U} \sum'_{\langle ij \rangle,\sigma} \left( n_{i,\sigma}(1 - n_{i,-\sigma})n_{j,-\sigma}(1 - n_{j,\sigma}) - n_{i,-\sigma}n_{i,\sigma}(1 - n_{j,-\sigma}) - n_{i,-\sigma}n_{j,-\sigma} + c_{i,\sigma}^\dagger c_{i,\sigma}^\dagger c_{j,\sigma} c_{j,-\sigma} \right)
\]

\[+ \frac{2t^2}{U} \sum_{\langle ij \rangle}(S_i \cdot S_j - \frac{1}{4} n_i n_j)
\]

(5.88)
If we now completely eliminate the doubly occupied states in the second and third terms, and assume that the charge-charge correlation effects are small, we would come to the final form of the effective Hamiltonian

\[
H = -t \sum_{(i,j),\sigma} (c_{i,\sigma}^\dagger c_{j,\sigma} + \text{H.C.}) + J \sum_{(i<j)} (S_i \cdot S_j - \frac{1}{4} n_i n_j) \tag{5.89}
\]

in which the coefficient \( J = 4t^2/U \), as the corresponding summation is running over the pair \((i, j)\) with \( i < j \) to avoid double counting. This is the same as the Heisenberg antiferromagnetic exchange with a hopping term, and is called the \( t-J \) model.

This model excludes doubly occupied states, and in the half-filling when the total number of carriers is equal to the total number of lattice sites, it reduces to the Heisenberg model for the antiferromagnetic insulators with \( s = 1/2 \).

### 5.2.3 The Schrieffer Wolff Transformation

In the third example, we will use the fact that the transformed operator \( \tilde{A} \) can be developed in a commutator expansion of the form as shown previously

\[
\tilde{A} = e^{-S} Ae^S = A + \frac{1}{1!}[A, S] + \frac{1}{2!}[[A, S], S] + \frac{1}{3!}[[[A, S], S], S] + \cdots \tag{5.90}
\]

Now, if we write \( H = H_0 + W \) where \( W \) is the interaction term, or non-diagonal term, the transformed Hamiltonian can be separated into two commutator series,

\[
\tilde{H} = H_0 + \frac{1}{1!}[H_0, S] + \frac{1}{2!}[[H_0, S], S] + \frac{1}{3!}[[[H_0, S], S], S] + \cdots + W + \frac{1}{1!}[W, S] + \frac{1}{2!}[[W, S], S] + \cdots \tag{5.91}
\]

The most simple requirement to eliminate the “first order” interaction, or non-diagonal terms is to set

\[
[H_0, S] = -W \tag{5.92}
\]

which reduces the original equation to

\[
\tilde{H} = H_0 + \left( \frac{1}{1!} - \frac{1}{2!} \right)[W, S] + \left( \frac{1}{2!} - \frac{1}{3!} \right)[W, S] + \left( \frac{1}{3!} - \frac{1}{4!} \right)[[W, S], S] + \cdots \tag{5.93}
\]

This method normally leads to a first order diagonalization which is adequate in many applications. In some cases, such as the Jahn-Teller problem [64], the transformation can lead to a diagonalization in the leading power of the coupling constant even in the strong coupling limit [86].

There are different ways to solve for \( S \) in Eq. (5.92) but the simplest way is to tackle it directly. In this example, we will transform the Anderson model using the
canonical transformation mentioned above. This transformation was in fact used by Schrieffer and Wolff in 1966 and is also commonly known as the Schrieffer and Wolff transformation.

Using the canonical transformation described in the last section for the periodic Anderson model Eq. (1.7), we can write the transformed Hamiltonian as

\[ \hat{H} = e^S H e^{-S} \]
\[ = H_0 + H_V + [S, H_0] + [S, H_V] + \frac{1}{2}[S, [S, H_0]] + \frac{1}{2}[S, [S, H_V]] + \cdots \] (5.94)

The variable \( S \) can be chosen in such a way to eliminate all terms linear in \( V_k \) as follows

\[ 0 = H_V + [S, H_0] \] (5.95)

and the transformed Hamiltonian becomes:

\[ \hat{H} = H_0 + \frac{1}{2}[S, H_V] + \frac{1}{3}[S, [S, H_V]] + \cdots \] (5.96)

The factor \( S \) satisfying Eq. (5.95) can be written as

\[ S = \sum_{k_\sigma} V_k \left( \frac{1 - n_{-\sigma}^f}{-\epsilon_k + \epsilon_f} + \frac{n_{-\sigma}^f}{-\epsilon_k + \epsilon_f + U} \right) (f_\sigma^\dagger c_{k_\sigma} e^{ikR_j} - c_{k_\sigma}^\dagger f_\sigma e^{-ikR_j}) \] (5.97)

\[ = \sum_{k_\sigma} V_k (A_k + Z_k f_{-\sigma}^f) (f_{-\sigma}^\dagger c_{k_\sigma} e^{ikR_j} - c_{k_\sigma}^\dagger f_{-\sigma} e^{-ikR_j}) \] (5.98)

where

\[ A_k = \frac{1}{-\epsilon_k + \epsilon_f} \] (5.99)
\[ Z_k = \frac{1}{-\epsilon_k + \epsilon_f + U} - \frac{1}{-\epsilon_k + \epsilon_f} \] (5.100)

It is easy to verify from evaluating the corresponding commutations, that this form of \( S \) satisfies Eq. (5.95). For instance, after expanding the commutations, one gets

\[ [S, H_0] = \sum_{jk_\sigma'} V_k ((A_k + Z_k n_{-\sigma'}^f)(\epsilon_k - \epsilon_f) - (A_k + Z_k)n_{-\sigma'}^f U) \cdot (f_{\sigma'}^\dagger c_{k_\sigma'} e^{ik'R_j} + c_{k_\sigma'}^\dagger f_{\sigma'} e^{-ik'R_j}) \] (5.101)

Using the definition of \( A_k \) and \( Z_k \) in Eq. (5.100)

\[ (A_k + Z_k n_{-\sigma'}^f)(\epsilon_k - \epsilon_f) - (A_k + Z_k)n_{-\sigma'}^f U = -1 \] (5.102)
and hence

\[
[S, H_0] = \sum_{j,k\sigma'} -V_k (f_{\sigma'}^\dagger c_{k'\sigma'} e^{i k' R_j} + c_{k'\sigma'}^\dagger f_{\sigma'} e^{-i k' R_j})
\] (5.103)

Having found \(S\), we can proceed to find the second term (first approximation) of the transformed Hamiltonian in Eq. (5.96), \(\frac{1}{2}[S, H_V]\), which is shown below:

\[
[S, H_V] = \sum_{i,j,k,k',\sigma,\sigma'} 2V_k^2 (A_k + Z_k n_{-\sigma'}) (f_{\sigma'} f_{\sigma'}^\dagger \delta_{k k'} \delta_{\sigma \sigma'} - c_{k\sigma}^\dagger c_{k'\sigma'} \delta_{ij} \delta_{\sigma \sigma'})
\]

\[\cdot e^{i(k' R_j - k R_i)} + Z_k (c_{k'\sigma'}^\dagger c_{k\sigma} f_{\sigma}^\dagger f_{-\sigma}^\dagger e^{-i(k' - k) R_j} + c_{-\sigma}^\dagger c_{k'\sigma'} f_{\sigma} f_{-\sigma} e^{i(k' - k) R_j})
\]

\[-Z_k (f_{-\sigma}^\dagger f_{\sigma}^\dagger c_{k\sigma} c_{k'\sigma'} e^{i(k + k') R_j} + f_{-\sigma}^\dagger f_{\sigma} c_{k-\sigma}^\dagger c_{k'\sigma'}^\dagger e^{-i(k + k') R_j})
\] (5.104)

This equation can be further simplified if we consider only those electrons whose momenta are close to the Fermi surface and assume that the on-site interaction is very strong. In other words, we approximate \(k' \approx k \approx k_F\) and ignore the phase factor in the last two terms. This simplifies the expression to:

\[
H_1 = \sum_{\sigma} \frac{V^2}{2} \left( 2(A + Z n_{-\sigma'}) (n_{\sigma'} - n_{k_F \sigma'}) + Z (2c_{k_F \sigma}^\dagger c_{k_F -\sigma} f_{-\sigma}^\dagger f_{\sigma}^\dagger - c_{k_F \sigma}^\dagger c_{k_F -\sigma} f_{\sigma}^\dagger f_{-\sigma}^\dagger + c_{k_F \sigma} c_{k_F -\sigma} f_{\sigma}^\dagger f_{-\sigma}^\dagger) \right)
\] (5.105)

where \(A\) and \(Z\) are \(A_k\) and \(Z_k\) with \(k\) substituted by \(k_F\). This result was first obtained by Schrieffer and Wolff in 1966, and has been used widely to relate the Anderson model to the Kondo model since then. The importance of this transformation lies in the new term of the transformed Hamiltonian:

\[
J_{k k'} = V^2 Z_k c_{k'\sigma'}^\dagger c_{k-\sigma} f_{\sigma}^\dagger f_{-\sigma}^\dagger
\] (5.106)

which provides a direct indication of spin interaction due to the hybridization \(H_V\) of the Hamiltonian.

Remarks on the results

The mathematics is straightforward canonical transformation, but the physics behind the mathematics is quite extraordinary. To start with, we first look at the condition in Eq. (5.95) on which the transformation is based. This equation is basically saying that commuting the non-interacting Hamiltonian \(H_0\) with \(S\) generates the negative of the hybridization term \(H_V\). Since the non-interacting Hamiltonian has no hybridization term, the commutation produces terms that are completely new to the original Hamiltonian. One way to look at it is to consider the canonical
transformation as a rotation operation in the Hilbert space, such as rotating an ellipse in a two dimensional plane. When an ellipse is formulated with its major axis on the x-axis and its minor axis on the y-axis, the \( x \) and \( y \) variables of the equation corresponding to this ellipse are separated without any cross term between the \( x \) and \( y \) variables. However, as soon as the ellipse is rotated to any angle other than the multiples of \( \pi/2 \), the corresponding equation will have a cross term which is a function of the angle of rotation. In our case, the non-interacting Hamiltonian has no cross terms and so the system is completely 'diagonalized'. It can also be considered as the superposition of two independent systems, one containing localized electrons and the other conduction electrons, without any mutual interactions. The canonical transformation in Eq. (5.94) however 'rotates' the system in such a way that the first order cross term of the resultant system is identical to the negative of the hybridization term \( H_V \), thereby cancelling the hybridization effect completely. This would be perfect if the transformation has only the first order terms. However, it is not the case here. The second order term will now dominate when the first order effect is cancelled, which contains the commutation of \( S \) with \( H_V \) as shown from Eq. (5.96). As the commutation of \( S \) with \( H_0 \) generates the hybridization term as mentioned before, the commutation of \( S \) with the hybridization term is bound to generate more new terms.

The result of this new commutation can be quite predictable if one thinks of the commutation with \( S \) as a hybridization generator. As we know, commutation of \( S \) with the non-interacting Hamiltonian generates the hybridized term \( H_V \). Commutation of \( S \) with the hybridized term would then hybridize the hybridized term, which yields products of electron operators with possibly different spins. This is exactly what the first order transformation result is. To be more precise, we first look at the expression of \( S \) in Eq. (5.98). It contains two terms, both in brackets. The first bracket can be written as the sum of \( (1 - n_{-\sigma}) \) and \( n_{-\sigma} \) with different coefficients, while the second bracket comprises the hybridization. Expanding the expression shows that \( S \) is the sum of two hybridizations, one with a singly occupied localized electron and the other without. If the parameter \( U \) is zero when the Coulomb interaction between localized electrons is not effective, the coefficients of \( (1 - n_{-\sigma}) \) and that of \( n_{-\sigma} \) will be the same, and the commutation result will be simplified to terms of \( f_{\sigma}^+ f_{\sigma} \) and \( c_{k\sigma}^+ c_{k\sigma} \). Hence, when double occupation has no impact on the Hamiltonian, the commutation with \( S \), or the hybridization generator, basically turns a localized electron and a conduction electron back to themselves. As a consequence, it contains the number operator of localized electrons and the potential scattering of conduction electrons.
5.2. Renowned Applications to Condensed Matter Systems

This potential scattering term was one of the major studies in the 50's, in an attempt to explain the effects of impurities inside a lattice of electrons with the introduction of an effective potential \( V \) due to the impurity. We know now that the theory alone is not sufficient to describe most transition metals and rare earth impurities, but it gives a good description of the virtual bound state of the conduction electrons around an impurity and the screening effect due to an impurity. For example, if an impurity has a charge in excess of that of the host metal atom, this will be a Coulomb potential. However no macroscopic electric field can exist within a metal as we know from classical field theory, the charge must be screened by the conduction electrons on an atomic scale. Both the Thomas-Fermi approximation and random phase approximation (RPA) can be used to estimate this screening effect, although the RPA has the advantage of including the long range oscillations of the screening cloud of the conduction electron (Friedel oscillations) which are not included in the Thomas-Fermi calculation. The starting Hamiltonian of the theories normally contains two terms, the kinetic energy of the conduction electrons and the one-body impurity potential scattering term.

\[
H = \sum_{k,\sigma} \epsilon_k c_{k\sigma}^\dagger c_{k\sigma} + \sum_{k, k', \sigma} V_{k, k'} c_{k\sigma}^\dagger c_{k'\sigma} \tag{5.107}
\]

in which the change of the conduction electron momentum in the second term is due to the Coulomb repulsion during the encounter with an impurity. This potential scattering modifies the density of states of the electrons near the impurity and induces extra states below the Fermi level. These extra states are created to accommodate the charge required to screen the impurity. If the local impurity potential \( V_{k, k'} \) is sufficiently attractive, an electron will be localized in the vicinity of the impurity, forming a bound state below the Fermi energy of the conduction electrons. On the contrary if the potential is not attractive enough to produce a bound state below the conduction band, it may nevertheless try to localize the conduction electrons for a period of time in the vicinity of the impurity. The state in which these conduction electrons are, is known as the virtual bound state. In such a state the conduction electrons spend a relatively longer period of time in the region of the impurity, but it is not a bound state as the wave function does not fall off exponentially with the distance from the impurity. Instead, it becomes the Bloch state of a free electron far from the impurity.

When \( U \) is not zero in the result of Eq. (5.105), the coefficient of \( n_{\uparrow, \sigma} \) will be smaller than that of \( (1 - n_{\downarrow, \sigma}) \) and creates an imbalance between terms of singly occupied states and terms of doubly occupied states. The result of the commutation with \( S \) in this case as shown in Eq. (5.105) contains extra terms that show spin
interaction and electron pair hybridization. Even with the terms that were in the result of \( U = 0 \), their coefficients in this case become more complex as the factor \( n^f_{-\sigma} \) in the term \( (1 - n^f_{-\sigma}) \) does not cancel out exactly the term \( n^f_{-\sigma} \) in the coefficient of Eq. (5.105). This results in a mixture of interactions. Part of the interactions would take place with the presence of an impurity while the other part would do without. There is however no spin interaction between electron operators in this interaction and so is ignored in most cases.

In a nutshell, we can summarize the result of the first order transformation as follows:

- \( c^\dagger_{k\sigma} c^\dagger_{k-\sigma} f^\sigma_{-\sigma} f^\sigma_{-\sigma} + \text{H.C.} \) is the pair hybridization term from which two conduction electrons are created on the expense of two localized electrons or vice versa. This is the result of two hybridizations applying on two similar electrons of different spin.

- \( c^\dagger_{k\sigma} c^\dagger_{k-\sigma} f^\dagger_{-\sigma} f^\sigma_{-\sigma} \) is the spin exchange term when the spin of the conduction electrons and the spin of the localized impurity are flipped at the same time. This is caused by two hybridizations applying on two different electrons of different spin. It is easier to see from its equivalent expression \( c^\dagger_{k\sigma} f^\sigma_{-\sigma} c^\dagger_{k-\sigma} f^\sigma_{-\sigma} \) that the result is the concatenation of two hybridizations.

- \( (A(1 - n^f_{-\sigma}) + Bn^f_{-\sigma}) n^f_{-\sigma} \) results from hybridizing an already hybridized localized electron. When \( U = 0 \), \( A = B \) and the \( n^f_{-\sigma} \) factor disappears, leaving a singly localized electron number operator. As the Coulomb interaction is absent, double occupation has no impact on the Hamiltonian and hence the result is independent of \( n^f_{-\sigma} \), the singly occupied localized electron of opposite spin. However when \( U \neq 0 \), the result becomes a mixture of localized electron number operators, some with single occupation while the other with double. This is the result of energy imbalance between single occupation and double occupation of an impurity site.

- \( (A(1 - n^f_{-\sigma}) + Bn^f_{-\sigma}) c^\dagger_{k\sigma} c^\dagger_{k'\sigma'} \) results from hybridizing an already hybridized conduction electron. As the momentum indices can be different between the two hybridizations, the conduction electron operators in this term do not result in number operators in general. Following the same reason above when \( U = 0 \) and \( A = B \), the result is independent of the presence of singly occupied localized electron of opposite spin. However, when \( U \neq 0 \), the result has a mixture of conduction electron interactions, some with single impurity occupation while others without. This is again the result of energy imbalance between single occupation and double occupation of a localized site.
According to this, commuting with $S$ is similar to applying a hybridization to the system and turns a conduction electron to a localized one or vice versa. As the transformation is the summation of results that are calculated from commuting the non-interacting system $H_0$ with $S$ $n$ times for $0 < n < \infty$, it resembles a perturbation of the Anderson model with respect to $V$. In fact, the factor $V^2$ in front of the first order transformation result does suggest that the result is to the lowest order of $V$. Also comparing this result with the result from chapter 4, the perturbative expansion in $V$ reveals that they are consistent with the lowest order of localized electron number operators.

Indeed the resemblance between the result from canonical transformation and the result from perturbative expansion with $V$ does make one to ponder the relationship between the two. This will be investigated further in the next chapters when the periodic Anderson model is considered in the canonical transformation for a more generic case. Higher order results are evaluated to confirm the relationship between the order of the transformation and the power of $V$ in the result.
We will show in this chapter the unitary transformation of the Periodic Anderson Model in one dimension, in a fashion similar to the Schrieffer and Wolff\textsuperscript{1} transformation of a single impurity Anderson Model in chapter 5. The resemblance of the single impurity Anderson Model to the Periodic Anderson Model is obvious when one compares the Hamiltonian of the two models in Eq. (1.7) and Eq. (4.1,4.2). It is therefore not surprising that the transformed Hamiltonians are very similar. The physics behind the two results are however quite different and this will be explained further in the chapter. In fact, the same principle can be applied to higher dimensions but the idea is best illustrated in one-dimension for its simplicity. The higher dimensional result will be presented in chapter 7.

The periodic Anderson model is believed to be one of the models best describing impurity effects and yet the understanding of it is still very limited. The model is solvable via Bethe Ansatz only in the single f state limit [88,89], i.e. for one f-electron. For the case of more then one electron, an exact solution is not possible. Even in the $U \to \infty$ limit when the periodic Anderson model becomes equivalent to the Kondo lattice model, an exact solution can only be obtained via approximations, such as bosonization [44]. There is, however, another approach to the understanding of the model whereby the Hamiltonian is transformed canonically to a form which shows the interaction more explicitly. This transformation is very similar to the one used by Schrieffer and Wolff in 1950 for the single impurity Anderson Model, as shown in chapter 5 in which the Hamiltonian of the Anderson Model was written in momentum space, with the impurity localized at the origin. An unitary transformation was applied to the Hamiltonian and the calculation was

\textsuperscript{1}As in the previous chapter, we shall always refer to the result of the transformation by Schrieffer and Wolff to Ref. [78] without explicit citing it.
performed up to the first order only. Even with this simple manipulation the transformed Hamiltonian shows explicit spin interactions between the localized impurity electron and the free conduction electrons, over a small range of parameters. This spin interaction is identical to the one that appears in the single impurity Kondo model, and as a result the transformation established a relationship between the single impurity Anderson model and the Kondo model. The drawback of the transformation is the limited range of parameters which confines the Anderson model to be Kondo like. It is basically due to the fact that the transformation was carried out only to the first order, assuming the higher order terms to become negligible in a restricted range of parameters. This problem can be overcome by calculating the transformation to higher orders, which in turn extends the validity range of the parameters. In fact, the transformation can be calculated up to infinite orders, with all the terms summed to a general expression. In the process of finding the general expression, a purpose built symbolic manipulation program was written to carry out the commutation of any fermionic operators using non-commutative algebra. The result of the commutations were then used in the canonical transformation. All of these will be the focus of this chapter, with the general result presented in the last section.

6.1 Canonical Transformation

The Hamiltonian of the periodic Anderson model considered here is the same as the one in Eqs. (4.1) and (4.2). The major difference of this Hamiltonian to that of Schrieffer and Wolff is merely the real and momentum space representations. The indices of the real space operators in the Hamiltonian can be associated with the site coordinate of electrons, which may provide information about the trajectories of electrons. This approach is especially effective and accurate if the electrons are moving slowly in the lattice as higher order hopping terms will be negligible and can be neglected for simplicity.

The unitary transformation we use as defined in chapter 5 is \( \tilde{H} = e^S H e^{-S} = H_0 + [S, Hv]/2 + [S, [S, Hv]]/3 + \ldots \) where \( S \) is determined by the condition \( Hv + [S, H_0] = 0 \). To satisfy this condition, for any value of \( U \) and \( V \), we use the standard Taylor expansion to linearize the atomic chain, i.e., \( c_{j \pm 1} \approx c_j \pm \partial c_j \) (the lattice constant was taken to be unity), and obtained:

\[
S = \sum_{i \sigma} V(A + Z f_{i,-\sigma} f_{i,-\sigma}) (f_{i,\sigma} c_{i,\sigma} - c_{i,\sigma} f_{i,\sigma}),
\]

where \( A = 1/(-2t + \mu_c + \epsilon_f) \) and \( Z = 1/(-2t + \mu_c + \epsilon_f + U) - A \). The parameter
A and \( Z \) are very similar to Eq. (5.95) in chapter 5 for the Schrieffer and Wolff transformation except that here it is in real space with an extra factor of \(-2t\). The linearization of the electron operator with respect to site distances is however new. It basically assumes that conduction electrons are moving with a constant velocity of \( \partial c_j \) in the lattice. In one dimension, this constant velocity is allowed only in two directions, namely left or right. If all electrons are moving in the same direction, the model would in fact be similar to a static model and possibly be approximated by one in the first order of Taylor expansion. However, if they are moving in a random direction, the momentum will be averaged out and the model is again to be approximated by a static model. This can indeed be verified by substituting the Taylor expansion of \( c_{i+1} \) in terms of \( c_i \) into the kinetic energy of the Hamiltonian.

Having found \( S \), one can obtain the first order transformation result (\( n = 1 \)).

\[
H_1 = \frac{1}{2} [S, HV] = \frac{V^2}{2} \sum_{j, \sigma} \left( 2(A + Z n^1_{j, -\sigma'})(n^f_{j, \sigma'} - n^c_{j, \sigma'}) \\
+ Z(2c^\dagger_{j, \sigma'} c_{j, -\sigma'} f^\dagger_{j, -\sigma'} f_{j, \sigma'} - c^\dagger_{j, \sigma'} c^\dagger_{j, -\sigma'} f^\dagger_{j, \sigma'} f_{j, -\sigma'} - c_{j, \sigma'} c_{j, -\sigma'} f^\dagger_{j, \sigma'} f^\dagger_{j, -\sigma'}) \right), \quad (6.2)
\]

The similarity of this result to that of Schrieffer and Wolff in Eq. (5.105) is apparent. The spin interaction, pair hybridization and potential scattering terms are all very much the same. The difference is merely the representation of electron operators. As no inter-site interaction among the electrons is present in the result, the first order transformation can be considered as a collection of independent systems, each of which behaves like a single impurity Anderson model at each site of the lattice. Another interesting point in the result is that the equation has only one site index, namely \( i \), which simplifies the result dramatically. This is however only true in one dimension when the movements of electrons are very restricted. It will be seen later in chapter 7 that the transformation in higher dimensions contains a summation of many different site indices.

Schrieffer and Wolff stopped the transformation at this stage, for the reason that higher order terms will be negligible under the condition: \( \gamma_0 \ll 1 \), or \( N(0)J_0 \ll 1 \). We, however, will continue to calculate the higher order terms. It is indeed the summation of these higher order terms that eventually gives us the exact coefficient of the spin interaction.

To evaluate the third Hamiltonian term \( \frac{1}{3} [S, [S, HV]] \), we require the result in Eq. (6.2) for \( H_1 \) from the last section and the commutation of it with \( S \). The third
6.1. Canonical Transformation

Hamiltonian term \( H_2 \), also named the second order transformation, is therefore:

\[
\frac{1}{3}[S, [S, HV]] = \sum_{i\alpha'} \frac{\nu^3}{3} \left( -4(A + Zn_{i-\sigma'}^f)^2 - 2Z(A + Zn_{i-\sigma'}^f)(n_{i-\sigma}^f - n_{i-\sigma'}^c) 
- 2Z^2(n_{i-\sigma'}^f - n_{i-\sigma'}^c)^2 \right) (c_{i\alpha'}^d f_{i\alpha'} + f_{i\alpha'}^d c_{i\alpha'})
\]  

(6.3)

Different products of the hybridization term with \( n_{i-\sigma'}^f \) and \( n_{i-\sigma}^c \) have returned from the algebra. In fact, we will find later that hybridization terms always reside in the even order Hamiltonian terms \( H_2, H_4, H_6, \ldots \).

It is also possible to rewrite the result in another format

\[
\left( A(1 - n_{i-\sigma}^f)(1 - n_{i-\sigma}^c) + B(1 - n_{i-\sigma}^f n_{i-\sigma}^c) + Cn_{i-\sigma}^f(1 - n_{i-\sigma}^c) + Dn_{i-\sigma}^f n_{i-\sigma}^c \right)
\cdot (c_{i\alpha'}^d f_{i\alpha'} + f_{i\alpha'}^d c_{i\alpha'})
\]  

(6.4)

where \( A, B, C \), and \( D \) are functions of \( A \) and \( Z \). This expression shows explicitly that the result contains a combination of four different hybridizations. The first one in the first factor is the hybridization with only a single electron. This electron can therefore hybridize without any hindrance of other electrons. The second term shows the second case in which hybridization occurs with a conduction electron always present. This would not cause any penalty to the system as double occupation of conduction electrons has no effect on the Hamiltonian. Similarly the third term shows another condition that the hybridization occurs with one localized electron always present. Contrary to the second case, this electron would have a significant impact on the Hamiltonian as the Coulomb interaction between localized electrons of the same site would create an energy barrier. Finally, the last term shows a hybridization with both conduction and localized electrons always present, which essentially is the combination of the last two cases.

Remembering that the transformation is formulated to remove the hybridization term in the Hamiltonian to the first order of \( V \) in Eq. (5.95), the same hybridization now re-appears in the third order of \( V \) with a more complex coefficient. In fact another transformation can be formulated to remove this generic hybridization term, but a more complex hybridization would re-appear in the fifth order anyway. After all, removing this hybridization term is not the focus of this chapter. We are more interested in the spin interaction of the Hamiltonian that shows up in the transformation which was implicit in the original formulation. To get a better understanding of this, going up to the higher order terms of the transformation is unavoidable.

We now turn to the next term of the transformation, the third order transformation which turns out to be more interesting than the last one. The spin interaction
6. One Dimensional Canonical Transformation

The coefficient of the spin interaction term versus $8V/U$ in the symmetric case at different order of the transformation, with $\mu_c = 0$.

Figure 6.1: The nominal coefficient of the spin-carrier interaction $J/V$ against $8V/U$ in the symmetric case at different order of the transformation.

The coefficient of the spin interaction term versus $8V/U$

$H_3 = \frac{1}{8} [\mathcal{S}, [\mathcal{S}, [\mathcal{S}, H_V]]] = \frac{V^4}{8} \sum_{j,\sigma} \left( 8A^3n_{j,\sigma}^c - 8A^3n_{j,\sigma}^f + (4AZ^2 - 4A^2Z)n_{j,\sigma}^c n_{j,\sigma}^f ight.
\left. - (8A^2Z + 8AZ^2 - 2Z^3)c_{j,\sigma}^c f_{j,\sigma'}^c f_{j,\sigma'}^f \right.
\left. + (16AZ^2 + 8Z^3)n_{j,\sigma}^f n_{j,\sigma}^c f_{j,\sigma'}^c \right.
\left. - (8AZ^2 + 8AZ^2 - 2Z^3)c_{j,\sigma}^c f_{j,\sigma'}^c f_{j,\sigma'}^f \right.
\left. + (16AZ^2 + 8Z^3)n_{j,\sigma}^f n_{j,\sigma}^c n_{j,\sigma}^c \right) (6.5)$

Compared with the coefficient of the spin interaction in $H_1$, the corresponding coefficient in this order has an extra factor of $-8V^2((A+Z)^2 + Z^2)$. Now if this extra
factor is less than minus one, the contribution of $H_3$ in the Hamiltonian regarding the spin interaction will be greater than that of $H_1$. This can also be verified from the curve $H_1 + H_3$ in figure 6.1 that this sum decreases to negative infinity as $Z$ increases (since $Z \propto \frac{1}{Z}$). The pair hybridization term also inherits a complex coefficient. In spite of all of these differences, it is still not difficult to find that this third order result is very similar to the first order result, albeit more complicated.

First of all, let us rewrite the third order result in Eq. (6.5) into another form

$$\frac{1}{8}[\mathcal{S}, [\mathcal{S}, [\mathcal{S}, H_V]]] = \frac{V^4}{8} \sum_{j\sigma,j'\sigma'} \left( (n_{j,\sigma'}^c - n_{j',\sigma'}^f)(8A^3 - (16AZ^2 + 8Z^3)n_{j,-\sigma}^c n_{j,-\sigma'}^c \right)$$

$$+ (4AZ^2 - 4A^2Z)n_{j,-\sigma}^c + (28A^2Z + 36AZ^2 + 16Z^3))n_{j,-\sigma'}^f$$

$$- (32A^2Z + 32AZ^2 + 16Z^3)c_{j,-\sigma}^j c_{j,-\sigma'}^f f_{j,-\sigma}^f f_{j,-\sigma'}$$

$$- (8A^2Z + 8AZ^2 - 2Z^3)(c_{j,-\sigma}^f c_{j,-\sigma'}^c f_{j,-\sigma}^f f_{j,-\sigma'}^f c_{j,-\sigma}^f c_{j,-\sigma'}^f f_{j,-\sigma}^f f_{j,-\sigma'}^f)$$

(6.6)

which can be categorized into three terms: a $(n_{j}\sigma - n_{j}\sigma')$ term, a spin interaction term and a pair hybridization term. The last two terms are identical to those of the first order result, except for the coefficients. The first term is however more generic in this order. It has factors that contain the four possible combinations of an electron in a single site, namely no electron occupation, one conduction electron, one localized electron and both conduction and localized electrons. Each configuration has its own coefficient in the result which shows the relative significance of the configuration in the system. Interestingly, all of these configurations have a common factor of $(n_{j}\sigma - n_{j}\sigma')$ which also appears in the first order result. A simple Hartree-Fock approximation could replace this by its average value in one of the terms and conclude that the significance of this term is proportional to the population of the conduction electron relative to that of the localized electron. This is however a very rough approximation as it would turn a two body interaction into a single body one.

A better way to do this is to rearrange these terms back to the original Hamiltonian $H_0$ and consider them as part of the unperturbed system. Moreover, bear in mind that it is only the third order result and it has already contained terms that are not in the lower order results. Higher transformation results may therefore have terms that are not present in the first three orders. It is therefore important to be able to calculate higher order terms of the transformation.

In a nutshell, the first and the second order results are completely orthogonal with no common terms but the first and the third order results seem to bear some similarities. These similarities may continue, as was initially speculated, to the higher order results. But without the actual results, it is very difficult to prove
6. One Dimensional Canonical Transformation

This. The next step is hence to calculate the higher order terms, to confirm this speculation and eventually look for a pattern based on these similarities.

The fourth order result is shown in the following:

\[
H_4 = \frac{1}{30} [\mathcal{S}, [\mathcal{S}, [\mathcal{S}, [\mathcal{S}, H_V]]]] = \frac{1}{30} \sum_{i\sigma} (16A^4 \\
+ (152A^3Z + 272A^2Z^2 + 196AZ^3 + 64Z^4)n_{i,-\sigma}^f \]

\[
-(88A^3Z + 88A^2Z^2 + 44AZ^3 - 4Z^4)n_{i,-\sigma}^c \]

\[
-(88A^2Z^2 + 88AZ^3 + 52Z^4)n_{i,-\sigma}^f n_{i,-\sigma}^c \]

\[
(c_{i,\sigma}^f f_{i,\sigma} + f_{i,\sigma} c_{i,\sigma}) \quad (6.7)
\]

and the fifth order result:

\[
H_5 = \frac{1}{144} [\mathcal{S}, [\mathcal{S}, [\mathcal{S}, [\mathcal{S}, H_V]]]] = \frac{1}{144} \sum_{i\sigma} \left( (512A^4Z + 1024A^3Z^2 + 1024A^2Z^3 + 512AZ^4 + 128Z^5) \\
+ (224A^4Z + 448A^3Z^2 + 328A^2Z^3 + 104AZ^4 - 4Z^5) \\
+ (336A^4Z + 848A^3Z^2 + 936A^2Z^3 + 520AZ^4 + 128Z^5)n_{i,\sigma}^f - n_{i,\sigma}^c \right) \\
+ (352A^3Z^2 + 528Z^2Z^3 + 368AZ^4 + 96Z^5)n_{i,-\sigma}^f n_{i,-\sigma}^c \]

\[
+ (176A^4Z + 176A^3Z^2 + 88A^2Z^3 - 8AZ^4)n_{i,-\sigma}^c n_{i,-\sigma}^f \quad (6.8)
\]

The calculation of these results are so complicated and tedious that they took days to work out by hand. The process is however very mechanical, based on the rules of the fermionic commutation. This kind of algebraic manipulation is in fact best suited for computers, for accuracy and repetitivity. A symbolic program was therefore written for this purpose, first to check the result of the first five orders, and then later to generate the results of higher orders. The first eleven order results generated from the program are shown in Appendix E.

A pattern seems to appear between the odd order results. A somewhat different pattern also appears in the even order commutation, which is completely different to the odd order one. Nevertheless, the odd and even order results are always related by a \( S \) commutation. This commutation determines the relationship of the coefficients between the two different orders.

Before proceeding further with the pattern, let us define the commutation of \( S \)
6.1. Canonical Transformation

with $H_V$ to be:

$$[[S, H_V]]_n = \underbrace{[S, [S, [S, \ldots, [S, H_V] \ldots]]}_{n \text{ times}},$$  \hspace{1cm} (6.9)

The pattern between the odd orders can then be put in the following form:

$$[[S, H_V]]_n = \sum_{i,\sigma} \left( J_n (c_i^\dagger c_{i,-\sigma} f_{i,-\sigma} f_{i,\sigma} - n_{i,\sigma} n_{i,-\sigma}^f) + P_n (c_i^\dagger c_{i,-\sigma} f_{i,-\sigma} f_{i,\sigma} + f_{i,\sigma} f_{i,-\sigma} c_i^\dagger c_{i,-\sigma}) + G_n (n_{i,\sigma}^e - n_{i,\sigma}^c) + I_n n_{i,-\sigma}^f n_{i,\sigma}^f + M_n n_{i,-\sigma}^e n_{i,-\sigma}^c (n_{i,\sigma}^f - n_{i,\sigma}^c) + K_n n_{i,-\sigma}^e n_{i,\sigma}^c \right),$$  \hspace{1cm} (6.10)

for odd values of $n$.

The value of the coefficients for the first order result are:

$$J_1 = 2ZV^2,$$
$$P_1 = -ZV^2,$$
$$G_1 = 2AV^2,$$
$$I_1 = 2ZV^2,$$
$$K_1 = 0,$$
$$M_1 = 0.$$

The same set of coefficients for the third order becomes:

$$J_3 = V^4(-32A^2Z - 32AZ^2 - 16Z^3),$$
$$P_3 = V^4(-8A^2Z - 8AZ^2 + 2Z^3),$$
$$G_3 = -8V^4A^3,$$
$$I_3 = V^4(-28A^2Z - 36AZ^2 - 16Z^3),$$
$$K_3 = V^4(-4A^2Z + 4AZ^2),$$
$$M_3 = V^4(16AZ^2 + 8Z^3).$$

In fifth order, it is:

$$J_5 = V^6(512A^4Z + 1024A^3Z^2 + 1024A^2Z^3 + 512AZ^4 + 128Z^5),$$
$$P_5 = V^6(224A^4Z + 448A^3Z^2 + 328A^2Z^3 + 104AZ^4 - 4Z^5),$$
$$G_5 = 32V^6A^5,$$
$$I_5 = V^6(336A^4Z + 848A^3Z^2 + 936A^2Z^3 + 520AZ^4 + 128Z^5),$$
$$K_5 = V^6(176A^4Z + 176A^3Z^2 + 88A^2Z^3 - 8AZ^4),$$
$$M_5 = V^6(-352A^3Z^2 - 528A^2Z^3 - 368AZ^4 - 96Z^5).$$
They all fall into the pattern formulated in Eq. (6.10). It is possible to keep finding the higher order results using the symbolic program to verify the pattern, but stating that Eq. (6.10) is true for any order of commutation is a completely different matter. To this end, an analytical tool is required to prove beyond doubt that Eq. (6.10) is true for any order. The best tool for this purpose is mathematical induction. In the following, we will write in detail every step of the proof, to ensure its accuracy. However when the algebra becomes too long and tedious, some of the non-essential steps will be skipped but without compromising on the accuracy of the context.

6.2 Proof by Induction

Before we start, two different indices are introduced here to differentiate the order of the commutation \( n \) from the recurrence of the coefficients \( J_m, P_m, \ldots \) over odd orders. The mapping of the two sequences can be written as \( n = 2m + 1 \) for odd order \( n \).

Knowing that Eq. (6.10) is true when \( n = 1 \) or \( m = 0 \), we assume Eq. (6.10) is also true for any \( m \). To find out whether the result is true for the next odd order, we first calculate its commutation with \( S \) for \([S, H_V]_{n+1}\), which yields:

\[
[S, H_V]_{n+1} = \sum_{i,\sigma} \left(-2V J_m(A + Zn_{i,-\sigma}^f) (r_{i,-\sigma}^f - n_{i,-\sigma}^c) 
-2VP_m(A + Zn_{i,-\sigma}^c) (n_{i,-\sigma}^f - n_{i,-\sigma}^c) - 2VG_m(A + Zn_{i,-\sigma}^f) 
-2VM_m(A + Z)n_{i,-\sigma}^f n_{i,-\sigma}^c - 2VI_m(A + Zn_{i,-\sigma}^f) n_{i,-\sigma}^f 
+2VK_m(A + Zn_{i,-\sigma}^f) n_{i,-\sigma}^c \right) \left(c^\dagger_{i,\sigma} f_{i,\sigma} + f^\dagger_{i,\sigma} c_{i,\sigma}\right). \tag{6.11}
\]

To show that Eq. (6.10) is true for any \( m \), Eq. (6.11) has to be commuted with \( S \) again. The algebra involved is long and tedious, and so only the result is reproduced here

\[
[S, H_V]_{n+2} = \sum_{i,\sigma} \left( J_{m+1}(c^\dagger_{i,\sigma} c_{i,-\sigma} f^\dagger_{i,-\sigma} f_{i,\sigma} - n_{i,-\sigma}^c n_{i,-\sigma}^f) 
+P_{m+1}(c^\dagger_{i,\sigma} c_{i,-\sigma} f_{i,\sigma} f^\dagger_{i,-\sigma} + f^\dagger_{i,\sigma} f_{i,-\sigma} c_{i,\sigma} c_{i,-\sigma}) + G_{m+1}(n_{i,\sigma}^f - n_{i,\sigma}^c) 
+I_{m+1} n_{i,-\sigma}^f n_{i,-\sigma}^c + K_{m+1} n_{i,-\sigma}^c n_{i,-\sigma}^c + M_{m+1} n_{i,-\sigma}^f n_{i,-\sigma}^c (n_{i,\sigma}^f - n_{i,\sigma}^c) \right), \tag{6.12}
\]
6.3. Evaluating the Coefficients

where

\[
\begin{align*}
J_{m+1} &= -J_m 4V^2((A + Z)^2 + A^2) - P_m 8V^2(A^2 + AZ) \\
&\quad - G_m 4V^2(2AZ + Z^2) - I_m 4V^2(A + Z)^2 - K_m 4V^2 A^2, \\
P_{m+1} &= -J_m 4V^2(A^2 + AZ) - P_m 2V^2((A + Z)^2 + A^2) \\
&\quad - (I_m + K_m) 2V^2(A + Z)A, \\
I_{m+1} &= -J_m 4V^2(A + Z)^2 - P_m 4V^2(A^2 + AZ) \\
&\quad - G_m 4V^2(2AZ + Z^2) - I_m 4V^2(A + Z)^2, \\
K_{m+1} &= -J_m 4V^2 A^2 - P_m 4V^2(A^2 + AZ) - K_m 4V^2 A^2, \\
M_{m+1} &= K_m 4V^2(2AZ + Z^2) - M_m 4V^2(A + Z)^2 \\
&\quad + J_m 4V^2(2AZ + Z^2), \\
G_{m+1} &= -G_m 4V^2 A^2.
\end{align*}
\]

Notice that Eq. (6.12) does not contain any new term that is not in \([S, H_V]_n\). By mathematical induction, we conclude that the form of the \(n\)th commutation of \(H_V\) with \(S\) is closed and is always given by Eq. (6.10).

### 6.3 Evaluating the Coefficients

The recursive Eq. (6.13-6.18) from the last section can be solved simultaneously to give the odd order coefficients of the transformed Hamiltonian, \(J_m, P_m, I_m, K_m\) and \(M_m\).

These recursive equations can be summarized into a matrix form in which

\[
\begin{pmatrix}
J_{m+1} \\
2P_{m+1} \\
I_{m+1} \\
K_{m+1}
\end{pmatrix}
= -4V^2 \mathbf{M} \cdot
\begin{pmatrix}
J_m \\
P_m \\
I_m \\
K_m
\end{pmatrix}
- 4V^2
\begin{pmatrix}
\alpha^2 - \beta^2 \\
0 \\
\alpha^2 - \beta^2 \\
0
\end{pmatrix}
\mathbf{G}_m,
\]

where \(\alpha = A + Z\) and \(\beta = A\). \(\mathbf{M}\) is a matrix given by:

\[
\mathbf{M} = \begin{pmatrix}
\alpha^2 + \beta^2 & 2\alpha\beta & \alpha^2 & \beta^2 \\
2\alpha\beta & \alpha^2 + \beta^2 & \alpha\beta & \alpha\beta \\
\alpha^2 & \alpha\beta & \alpha^2 & 0 \\
\beta^2 & \alpha\beta & 0 & \beta^2
\end{pmatrix}
\]

(6.20)

Not only is this matrix \(\mathbf{M}\) symmetric, but its determinant is zero. It is, in fact, easy to see this since the first row of the matrix is equal to the sum of the third and
the fourth row, while the second row can be expressed as the linear combination of the last two rows. Explicitly,

\[ J_{m+1} = I_{m+1} + K_{m+1} , \]  
\[ 2P_{m+1} = \frac{\beta}{\alpha} (I_{m+1} + G_{m+1}) + \frac{\alpha}{\beta} (K_{m+1} - G_{m+1}) . \]  

The expression of \( I_{m+1} \) and \( K_{m+1} \) in terms of \( J_{m+1} \) and \( P_{m+1} \) can be deduced directly from Eq. (6.19), together with the general solution of \( G_m \).

\[ I_{m+1} = \frac{\alpha \beta}{\beta^2 - \alpha^2} (2P_{m+1} - \frac{\alpha}{\beta} J_{m+1}) - G_{m+1} , \]  
\[ K_{m+1} = \frac{\alpha \beta}{\alpha^2 - \beta^2} (2P_{m+1} - \frac{\beta}{\alpha} J_{m+1}) + G_{m+1} , \]  
\[ G_m = (-2V^2)^m (2\beta \alpha^2)^m G_0 . \]

Substituting these equations back into Eqs. (6.13) and (6.14) yields the surprisingly simple result:

\[ J_{m+1} = -2^3 V^2 (\alpha^2 + \beta^2) J_m , \]  
\[ P_{m+1} = -2V^2 (3 \alpha \beta J_m + (\alpha^2 + \beta^2) P_m) . \]

In terms of \( J_0 \) and \( P_0 \) they can be written as:

\[ J_m = (-2^3 V^2)^m (\alpha^2 + \beta^2)^m J_0 , \]  
\[ P_m = (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} (\alpha^2 + \beta^2) P_0 + \alpha \beta (4^m - 1) J_0 . \]

Using Eqs. (6.21) and (6.22), \( K_{m+1} \) can be written in terms of \( P_m, J_m, G_m \) and subsequently be solved by using Eqs. (6.25, 6.28, 6.29):

\[ K_m = (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} \left[ (4^m \beta^2 - \frac{2 \alpha^2 \beta^2}{\alpha^2 - \beta^2}) J_0 + 2 \alpha \beta \frac{\alpha^2 + \beta^2}{\alpha^2 - \beta^2} P_0 \right] + G_m . \]

\[ I_m \] can now be evaluated from \( K_m \) using Eq. (6.21):

\[ I_m = (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} \left[ (4^m \alpha^2 + \frac{2 \alpha^2 \beta^2}{\alpha^2 - \beta^2}) J_0 - 2 \alpha \beta \frac{\alpha^2 + \beta^2}{\alpha^2 - \beta^2} P_0 \right] - G_m . \]

\( M_m \) is slightly more complicated since it depends on not only \( M_{m-1} \), but also \( K_{m-1} \) and \( J_{m-1} \). However, it can still be summed after some algebraic manipulations:

\[ M_m = -(-2^3 V^2)^m (\alpha^2 + \beta^2)^{m-1} (\alpha^2 - \beta^2) J_0 \]
\[-[-2V^2(\alpha^2 + \beta^2)]^m \frac{4 \alpha \beta}{\alpha^2 - \beta^2} \left( \frac{\alpha \beta J_0}{\alpha^2 + \beta^2} - P_0 \right) \]
\[ +(-4V^2 \alpha^2)^m \left( \frac{\alpha^2 + \beta^2}{\alpha^2 - \beta^2} J_0 - \frac{4 \alpha \beta}{\alpha^2 - \beta^2} P_0 \right) \]
\[ -G_0 + M_0 + (-4V^2 \beta^2)^m G_0 . \]
These expressions can be further simplified by using the first order results, which gives

\[ J_m = (-2^3V^2)^m (\alpha^2 + \beta^2)^m 2(\alpha - \beta)V^2, \quad (6.33) \]
\[ P_m = (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} [2^{m+1}\alpha\beta - (\alpha + \beta)^2] (\alpha - \beta)V^2, \quad (6.34) \]
\[ G_m = (-2V^2)^m (2\beta^2)^m 2\beta V^2, \quad (6.35) \]
\[ K_m = \left\{ (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} [4^m\beta(\alpha - \beta) - \alpha(\alpha + \beta)] + (-2V^2)^m (2\beta^2)^m \right\} 2\beta V^2, \quad (6.36) \]
\[ I_m = (-2^3V^2)^m (\alpha^2 + \beta^2)^{m-1} 2\alpha^2(\alpha - \beta)V^2 \]
\[ + (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} 2\alpha\beta(\alpha + \beta)V^2 \]
\[ - (-2V^2)^m (2\beta^2)^m 2\beta V^2, \quad (6.37) \]
\[ M_m = -(-2^3V^2)^m (\alpha^2 + \beta^2)^{m-1} 2(\alpha - \beta)(\alpha^2 - \beta^2)V^2 \]
\[ + (-2V^2)^m (\alpha^2 + \beta^2)^{m-1} 4\alpha\beta(\alpha + \beta)V^2 \]
\[ + (-4\alpha^2V^2)^m 2\alpha V^2 + (-4\beta^2V^2)^{m} 2\beta V^2. \quad (6.38) \]

The even order coefficients can be deduced from the odd order coefficients, after rearranging the \( n + 1 \)th commutation result Eq. (6.11) to the form:

\[ [[S, \mathcal{H} V]]_{n+1} = \sum_{i,\sigma} \left( R_m + S_m n_{i,\sigma}^f + T_m n_{i,-\sigma}^c + Q_m n_{i,-\sigma}^f n_{i,\sigma}^c \right) (c_{i,\sigma}^f f_{i,\sigma}^f + f_{i,\sigma}^f c_{i,\sigma}^f), \quad (6.39) \]

where

\[ R_m = -2VG_m A, \]
\[ S_m = -2V[(J_m + P_m)A + I_m(A + Z) + (J_m + G_m)Z], \]
\[ T_m = 2V[(J_m + P_m + K_m)A + P_m Z], \]
\[ Q_m = 2V[(J_m + K_m - P_m)Z - M_m(A + Z)]. \quad (6.40) \]

The general solutions of \( J_m, P_m, G_m, I_m, K_m, M_m \) from Eqs. (6.33-6.38) are used to find the expression of the even order coefficients:

\[ R_m = -(-4V^2\beta^2)^m 4\beta^2 V^3, \quad (6.41) \]
\[ S_m = -[-2V^2(\alpha^2 + \beta^2)]^m 8\alpha(\alpha - \beta)V^3 \]
\[ + (-2V^2(\alpha^2 + \beta^2))^m 2\beta(\alpha + \beta)V^3 \]
\[ + (-4V^2\beta^2)^m 4\beta^2 V^3, \quad (6.42) \]
6.4 The Transformed Hamiltonian

Using these general expressions for the \( n \)th and the \( n+1 \)th commutation of \( S \) with \( H_V \), the exact infinite order transformation of the Hamiltonian can be written down. The transformed Hamiltonian comprises \( H_0 \), the sum of the odd order commutations of \( S \) with \( H_V \) and the sum of the even order commutations:

\[
\tilde{H} = H_0 + H_{\text{odd}} + H_{\text{even}},
\]

where

\[
H_{\text{odd}} = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] [[S, H_V]]_{2m+1},
\]

\[
H_{\text{even}} = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+2)!} - \frac{1}{(2m+3)!} \right] [[S, H_V]]_{2m+2}.
\]

It is straightforward to sum the odd order Hamiltonian to get:

\[
H_{\text{odd}} = \sum_{i,\sigma} J(c_{i,\sigma}^\dagger c_{i,-\sigma} f_{i,\sigma}^\dagger f_{i,-\sigma} - n_{i,\sigma}^c n_{i,-\sigma}^c) + P(c_{i,\sigma}^\dagger c_{i,-\sigma} f_{i,\sigma}^\dagger f_{i,-\sigma} + f_{i,\sigma}^\dagger f_{i,-\sigma} c_{i,\sigma} c_{i,-\sigma}) + G(n_{i,\sigma}^f - n_{i,\sigma}^c) + M n_{i,-\sigma}^f n_{i,\sigma}^c (n_{i,\sigma}^f - n_{i,\sigma}^c) + K n_{i,-\sigma}^c n_{i,\sigma}^c.
\]

where \( J, P, G, I, K \) and \( M \) are the summation of the corresponding coefficients \( J_m, P_m, G_m, I_m, K_m \) and \( M_m \) over an infinite number of \( m \). If we define \( \theta = \sqrt{2V^2(\alpha^2 + \beta^2)}, \theta_\beta = 2V\beta, \theta_\alpha = 2V\alpha \) and \( F(x) = \sin x/x + (\cos x - 1)/x^2 \) then the exact values of the coupling constants from Eq. (6.48) are:

\[
J = 2(\alpha - \beta)V^2 F(2\theta),
\]

\[
P = 2\alpha\beta(\alpha - \beta)V^2 F(2\theta) - (\alpha - \beta)V^2 \frac{(\alpha + \beta)^2}{\alpha^2 + \beta^2} F(\theta),
\]

\[
G = 2\beta^2 V^2 F(\theta_\beta),
\]

\[
K = 2\beta^2 V^2 \frac{\alpha - \beta}{\alpha^2 + \beta^2} F(2\theta) - 2\alpha\beta V^2 \frac{\alpha + \beta}{\alpha^2 + \beta^2} F(\theta) + G,
\]
6.4. The Transformed Hamiltonian

\[
I = 2\alpha^2 V^2 \frac{\alpha - \beta}{\alpha^2 + \beta^2} F(2\theta) + 2\alpha\beta V^2 \frac{\alpha + \beta}{\alpha^2 + \beta^2} F(\theta) - G, \quad (6.53)
\]
\[
M = -2(\alpha - \beta) \frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2} V^2 F(2\theta) - 4\alpha\beta \frac{\alpha + \beta}{\alpha^2 + \beta^2} V^2 F(\theta) + 2\alpha V^2 F(\theta_\alpha) + 2\beta V^2 F(\theta_\beta). \quad (6.54)
\]

Similarly, the even order Hamiltonian can be evaluated by substitution

\[
H_{even} = \sum_{i,\sigma} (R + S n_i^\dagger + T n_i + Q n_i^\dagger n_i^\dagger) (c_{i,\sigma} f_{i,\sigma} + f_{i,\sigma}^\dagger c_{i,\sigma}), \quad (6.55)
\]

where \( R, S, T \) and \( Q \) are the summation of the corresponding \( R_m, S_m, Q_m \) and \( T_m \) over infinite number of \( m \). Using the same notations \( \theta, \theta_\beta, \theta_\alpha \) as in the odd order coefficients, and \( F'(x) = \sin x/x^3 - \cos x/x^2 \), we obtain:

\[
R = -4\beta^2 V^3 F'(\theta_\beta), \quad (6.56)
\]
\[
S = -8(\alpha - \beta)\alpha V^3 F'(2\theta) - 2(\alpha + \beta)\beta V^3 F'(\theta) - R, \quad (6.57)
\]
\[
T = 8\beta(\alpha - \beta) V^3 F'(2\theta) - 2\alpha(\alpha + \beta) V^3 F'(\theta) - R, \quad (6.58)
\]
\[
Q = 8(\alpha - \beta)^2 V^3 F'(2\theta) + 2(\alpha + \beta)^2 V^3 F'(\theta)
-4\beta V^3 F'(\theta_\beta) - 4\alpha V^3 F'(\theta_\alpha). \quad (6.59)
\]

For the symmetric case \( \epsilon_f = U/2 \), we have \( Z = -2A, \theta_\beta = -\theta_\alpha = 2AV, \theta = |2AV|, \alpha = -A \) and \( \beta = A \), both the odd and even order Hamiltonian coefficients simplify considerably. \( A = -2/U \) and \( Z = 4/U \). \( \theta \) is assumed to be positive but \( 2AV \) in this case is most likely to be negative. Since \( F(x) \) is an even function, it makes no difference if we write \( F(|x|) \) or \( F(x) \) and so we will choose the latter case for simplicity. The odd order becomes

\[
J_{sym} = -4AV^2 F(4AV), \quad (6.60)
\]
\[
P_{sym} = 2AV^2 F(4AV) = -\frac{1}{2} J_{sym}, \quad (6.61)
\]
\[
G_{sym} = 2AV^2 F(2AV), \quad (6.62)
\]
\[
K_{sym} = -2AV^2 (F(4AV) - F(2AV)), \quad (6.63)
\]
\[
I_{sym} = -2AV^2 (F(4AV) - F(2AV)), \quad (6.64)
\]
\[
M_{sym} = 0. \quad (6.65)
\]

and the even order is

\[
R_{sym} = -4A^2 V^3 F'(2AV), \quad (6.66)
\]
\[
S_{sym} = -16A^2 V^3 F'(4AV) - R_{sym}, \quad (6.67)
\]
\[
T_{sym} = -16A^2 V^3 F'(4AV) - R_{sym}, \quad (6.68)
\]
\[
Q_{sym} = 32A^2 V^3 F'(4AV) + 2R_{sym}. \quad (6.69)
\]
These results are very different to what one would expect from the model, since the variations due to the sinusoid functions were not foreseeable. In the first order result from Schrieffer and Wolff, this kind of behaviour did not appear and no one would have anticipated a phenomenon so different to that of the first order would surface in the higher orders. To understand the implications of the results and its significance, we will first look at the symmetric case where the number of independent variables are minimal. The generic asymmetric case will be considered afterwards.

6.4.1 Symmetric Case

In the symmetric case when \( \epsilon_f = -U/2 \), the number of independent variables is reduced to one, in which \( t \) and \( \mu_c \) are assumed constant as usual. The spin-spin interaction described from the coefficient \( J \) becomes very simple, as shown in figure 6.2. The dashed line in the figure shows the first order result of Schrieffer Wolff
while the solid line indicates our sum of infinite order result. The variation of $J$ with $U$ and $V$ is quite surprising, from which an attractive spin carrier interaction is implied when $J$ falls below zero. This however does not happen for all $U$ and $V$ and the system has to be set to a particular condition before $J$ becomes negative. Looking from another angle, the system has a phase diagram with at least two different phases acting together, namely ferromagnetic and antiferromagnetic. When $J$ is positive, electrons of opposite spin tend to stay together forming a pair to lower the energy of the system. This as a consequence drives the system to an antiferromagnetic state. If $J$ is negative, electrons of the same spin tend to stay together which reduces the energy of the system and therefore forces the system to become ferromagnetic. The cross-over points of $J$ from positive to negative or vice versa, also indicate a state in which the system could be paramagnetic when the orientation of the spin makes no difference to the energy of the state.

This result is not unique to us. E. Müller-Hartmann and A. Reischl [69] showed indirectly a similar variational behaviour of the system at higher orders in three
dimensional space, which will be discussed in the next chapter.

The insert of figures 6.2 - 6.9 shows the corresponding plot with the x-axis inverted. In this way, we can see the oscillation continues to infinite $V$ or zero $U$ with only a slight change of magnitude. The value $J$ is however still undetermined when $U$ is zero, and therefore remains as a singular point.

It is clear from the figures of all other coefficients (6.4-6.6), that Schrieffer and Wolff's result has a singular point at the origin for all coefficients, except those that are identically zero. These singular points are all of first order, due to the fact that the result of Schrieffer and Wolff is only up to the first order. With our sum to infinite orders, the singular points still remain in the result and are located at the same place, but are of infinite order. This can be seen from the solid lines in the figures that converge to zero sinusoidally with increasing frequency but the magnitude is fairly constant. The cause of the singular point can be traced back to the parameter $S$ in the canonical transformation of Eq. (5.6) which prohibits $U$ to be zero. In other words, a pole at $U = 0$ is added to the Hamiltonian with every com-

Figure 6.4: Coefficient $G$ vs $U/8V$ in the symmetric case from our canonical transformation (solid line) and Schrieffer Wolff's (dashed line) with $t = 0, \mu_c = 0$. The insert shows the same plot with x-axis inverted.
mutation with $S$. Since the $n$th order result of the transformation contains $n$ times commutation of $H_1$ with $S$, there must be $n$ poles in the result at the same place for $n$th order transformation. Summing the infinite order series therefore produces a singular point of infinite order. The singular point is nonetheless unavoidable if canonical transformation of this kind is used, unless the parameters of the transformation are chosen in such a way that the denominator will never be zero. This will be discussed in chapter 7 in more detail. In fact, the ripple (variation) of the coefficient looks as if it was created by the suppression of the first order singular point from higher order terms. The more the higher order terms are considered in the transformation, the harder the singularity is suppressed and more ripples resulted.

When $U$ is large compared with $V$, our result approaches the result of Schrieffer and Wolff as indicated by the convergence of the solid line (our result) to the dashed line (Schrieffer and Wolff’s result) in the figures. Since the convergence occurs at large $U/V$, it appears on the left of the main figures and on the right of the inserts. This can also be confirmed from the corresponding Eqs. (6.60 - 6.64). All of these
6. One Dimensional Canonical Transformation

![Graph showing coefficients](image)

**Figure 6.6:** Coefficient $K$ vs $U/8V$ in the symmetric case from our canonical transformation (solid line) and Schrieffer Wolff’s (dashed line) with $t = 0, \mu_c = 0$. The insert shows the same plot with x-axis inverted.

in fact show that the first order term of the transformation is dominant when $U$ is large. In other words, when the penalty of double occupation is so large that only electrons with a substantial amount of energy would be able to doubly occupy a site, the probabilities of hybridization and hopping are greatly reduced. The higher order interactions of electrons become rather insignificant compared with the first order result, and hence the first order term of the transformation is adequate to describe the behaviour of the system.

One of the interesting points of the symmetric case is that it reduces the parameter space of the model significantly. For instance, we can summarize the coefficients...
in which there are only two independent variables, namely $J_{\text{sym}}$ and $G_{\text{sym}}$. Because $J_{\text{sym}}$ varies twice as fast as $G_{\text{sym}}$ does, the summation and subtraction of the two create different behaviours for coefficients $K_{\text{sym}}$ and $I_{\text{sym}}$ as shown in figure 6.5 and figure 6.6. These two coefficients $K_{\text{sym}}$ and $I_{\text{sym}}$ are in fact quite similar when plotted with $8V/U$ as shown in the inserts of the figures 6.5 and 6.6.

The coefficient $M$ is identically zero in the symmetric case for both our result and the result of Schrieffer and Wolff. They are hence not shown.

The even order coefficients of the transformation result also contain a similar
oscillation. The variation shown in figure 6.7 - 6.9 looks very similar to the odd order coefficients, as a result of the combinations of two functions with harmonically related frequency content in the sine and cosine functions. The singularity is of infinite order and remains staying at the origin as in the odd order terms. Since Schrieffer and Wolff did not calculate the second order transformation which is equivalent to our first even order, our figures have nothing to compare against and so only solid lines are shown for all even order coefficients.

Similar to the odd order terms, the parameter space of the model in the even order terms is greatly reduced for the symmetric case. The coefficients in Eqs. (6.66 - 6.69) can be rewritten as:

\[
R_{\text{sym}} = -4A^2V^3F'(2AV), \quad (6.75) \\
S_{\text{sym}} = -16A^2V^3F'(4AV) + R_{\text{sym}}, \quad (6.76) \\
T_{\text{sym}} = R_{\text{sym}}, \quad (6.77) \\
Q_{\text{sym}} = -2S_{\text{sym}} + 4R_{\text{sym}}. \quad (6.78)
\]

in which there are only two independent variables \( R_{\text{sym}} \) and \( S_{\text{sym}} \).
6.4. The Transformed Hamiltonian

Many of the terms we discussed above from the transformation are new, but some of them already exist in the unperturbed Hamiltonian $H_0$ of Eq. (4.1). These include $Gn_i^{f\sigma}$, $-Gn_i^{c\sigma}$ and $+In_{i,-\sigma}^{f}n_i^{f\sigma}$ which should renormalize the corresponding terms in $H_0$. Figure 6.10 shows the coefficients of $n_i^{f\sigma}$ with respect to $U/8V$ after renormalization. Since $\epsilon_f = -U/2$ in the symmetric case, the unperturbed Hamiltonian $H_0$ merely adds $-U/2V$ to the coefficient $G/V$ for renormalization. A very similar renormalization also happens to $n_i^{c\sigma}$ which changes the coefficients from $-G$ to $-G - \mu_c$. The $\mu_c$ is contributed from $H_0$ and is assumed zero. This coefficient looks exactly like $G$ in figure 6.4. The last term altered by the renormalization is the Coulomb interaction $n_{i,-\sigma}^{f}n_i^{f\sigma}$, to which a $U$ is added due to the contributions from $H_0$. The renormalized Coulomb interaction is plotted in figure 6.11, with very similar behaviour as the previously renormalized coefficient.

In summary, oscillation appears in all coefficients, except possibly $M$ which is identically zero. Even though the frequency of the variation varies from one coefficient to another, the patterns are nevertheless very similar. Our results always converge to the results of Schrieffer and Wolff at large $U$ or small $V$, although the figure might not be able to show it for a limited range of $U/8V$. The singularity of our results is always of higher orders than that of Schrieffer and Wolff as higher
order terms are included in our transformation. In the case of coefficient $K$, we have a singularity at $U = 0$, when Schrieffer and Wolff predicted $K$ to be identically zero. This is the obvious evidence that the result of the first order transformation is not enough to provide a full picture of the system.

6.4.2 Asymmetric Case

When $\epsilon_f \neq -U/2$, the energy levels of the singly occupied and doubly occupied localized electron states are no longer symmetrically sitting on the upper and the lower sides of the conduction energy band. Therefore, the energy required to doubly occupy a localized site bears no relationship to the single occupation energy level. As a result, the number of independent parameters increases and the parameter space extends beyond that of the symmetric case.

Even with this extended parameter space, the behaviour of the coefficients still resembles that of the symmetric case. One particular example of the asymmetric
6.4. The Transformed Hamiltonian

Figure 6.11: Renormalized Coefficient of $n^+_\uparrow n^\uparrow_\downarrow$ vs $U/8V$ in the symmetric case from our canonical transformation (solid line) and the Schrieffer and Wolff's (dashed line) with $t = 0, \mu_c = 0$. The insert contains the same curves with x-axis inverted.

The transformed Hamiltonian

The transformed Hamiltonian when $V$ is 1 and $U$ is 8 is shown in figure H.1 to H.29 in Appendix H. The difference of the asymmetric case to the symmetric case is merely the form of the variation and position of the singular point. The general behaviour of the model is otherwise the same. It is more obvious in these figures that our result approaches the result of Schrieffer and Wolff for large $U$ and small $V$, when the grey long dashed line and the grey dotted line join together at small value of $V$ and converge at large $U$.

One interesting point in most of the figures in the asymmetric case is that the number of maximum singular points in the system has increased to two, situated at $\epsilon_F = 0$ and $\epsilon_F = -U$. These singularities again are the artefact of the transformation as the model itself does not become singular at these conditions. However, as long as the system is not at any of these two singular points the transformation is still exact and the result still valid. The oscillation is a genuine effect of the transformation and is the consequence of the transformation applied on the model in the Hilbert Space. Other transformations exist to remove the oscillations from the Hamiltonian, as described in chapter 7, but they are in general different.
transformations. As far as this transformation is concern, the oscillation is a basic characteristic of the result which is always presence near the singularities.

Among the plots, the vertical scales are different and so re-scaling is needed for comparison. The singular point is sitting at $U = 3$ in many cases where oscillation appears to be most significant. In some other plots such as $G$ and $R$, the coefficients turn out to be independent of $U$ and hence the singularity is absent. The variation of the coefficients against $V$ is sinusoidal, except for $S$ and $T$ where they appear to be more irregular. This is due to the combination of higher order sine and cosine functions included in the result. More discussion of the asymmetric case can be found in Appendix H.

6.5 Remarks

The oscillation in the result of the transformation is believed to come from two competing effects originating from the alternating odd order terms of the transformation. Because successive odd order terms are obtained from commutation with $S$, the oscillation must be a consequence of the form of $S$ chosen. As $S$ is chosen to produce the exact first order hybridization term of opposite sign when commuted with $H_V$, it produces hybridization terms $H_V$ of opposite sign whenever it is commuted with terms similar to $H_0$. When the higher order terms of the transformation are in fact similar to $H_V$ for every second order, the sign of the result alternates for every second commutations. This explains why the results from two successive odd order terms are competing or cancelling.

Based on the same principle, it is also possible to transform the Hamiltonian to a form free of oscillations as mentioned previously. The details of this are covered in the next chapter, in which three dimensional Anderson models will be considered.
CHAPTER 7

The Three Dimensional Periodic Anderson Model

In chapter 6, we have seen how canonical transformation is introduced to the periodic Anderson model in a one dimensional space and the creation of singular points as a result of the transformation. These singular points are inherited from the poles of the coefficients in $S$ of Eq. (6.1) in the transformation and therefore exist in all orders of the results. Since each order introduces an extra pole to the result from the $S$ in the commutation, the singularity created in the $n$th order term is therefore of $n$th order. This high order singularity turns a simple first order pole in the result of Schrieffer and Wolff into high order ripples as shown in figure 6.2 for instance. The frequency of the ripple increases as the function approaches the singular point. This already has several implications to the physical behaviour of a one dimensional model. It will be interesting to find out how the model behaves in three dimensions.

A first order expansion of the electron operator with respect to continuous interatomic distance was used in the one dimensional case to simplify the formulation and make the transformation possible. This kind of expansion is commonly used in the bosonization of one dimensional model, as the motion of electrons in one dimension is very restricted. In higher dimensions, this assumption would fail miserably and other alternatives are needed to simplify the calculation instead. Furthermore, the number of indices for each operator would increase dramatically in higher dimensions, due to the increased number of degrees of freedom available for each electron. The complexity on the behaviour of electrons would as a result increase exponentially. Some collective phenomena could also arise in the process, which is our main motivation for the calculation.

1As in the previous chapter, we shall always refer to the result of the transformation by Schrieffer and Wolff to Ref. [78] without explicit citing it.
To study the three dimensional system using our existing tools, we need to extend the same transformation we used from chapter 6 to a three dimensional model and impose different conditions to simplify the calculation. In general, conduction electrons can hop from site to site and turn into (hybridize) localized electrons. This makes the localized electrons appear as if they are moving as well. However, the strongest interaction must come from a close range and so the effective interaction must be dominant between electrons of nearest neighbours or of the same site. According to this, we managed to find the interaction of electrons to the lowest order using a transformation similar to the one in chapter 6. Our results are discussed with works of others, to compare the differences between the various approaches.

The benefit of using canonical transformation in our approach, as opposed to many other methods, comes from the fact that the transformation is unitary which conserves the eigenvalues of the system. The characteristics of models not so obvious from the form of the Hamiltonian will in many cases become apparent after the transformation. It is particularly useful in the Anderson Model for its vast parameter space with both spin and charge interactions. For instance, Schrieffer and Wolff's result only brought out the spin interacting sector of the model by eliminating the hybridization to the first order of $V$, as shown in chapter 6. The charge fluctuation is however still left behind in the higher order terms of the result. After all, the hybridization is removed only to the first order of $V$ which reduces but does not entirely remove the mobility of the localized $f$-electron that relies on the higher order hybridization to superexchange. This non-diagonal nature of the transformation causes the spin-interaction to stand out, and still allows electrons to hybridize in the background. Understanding the intricate interplay of the electrons in this case is difficult, without knowing the higher order terms of the transformation. One of the goals of this chapter is therefore to calculate the higher order terms and find out whether these terms have any influence on the physics of the model. If they do, the impact of these terms on the behaviour of the model will be of primary interest to us. Charge fluctuation and spin fluctuation could in fact be present in the result of the Schrieffer and Wolff transformation, but without going to higher order expansion we would not be able to see it. This chapter will show that spin and charge fluctuations indeed exist with respect to the parameters $U$ and $V$ of the model, from the Schrieffer and Wolff transformation.

The charge and spin fluctuations from the transformation seem to be related and are believed to be the consequence of higher order hybridization processes remaining in the transformed Hamiltonian. Using a slightly different transformation, we managed to remove the hybridization from the Hamiltonian and show that the
singularities disappear completely from the coefficients.

In the following, we will first introduce the Anderson Model in three dimensional space and the conditions under which the canonical transformation is applied. The mathematics of the transformation is very similar to that of chapter 6 and is therefore skipped to avoid repetition. The spin interaction from the result of the transformation is discussed in the context, which includes comparison with other results from the literature. Other transformations will also be discussed on removing the hybridization completely out of the Hamiltonian which shows very promising results.

7.1 Canonical Transformation

The Hamiltonian of the most general periodic Anderson model is provided in Eqs. 4.1 and 4.2

The restriction that conduction electrons and localized electrons in the hybridization term have to be at the same site as in the Anderson Model, is relaxed in this section as the most general configuration is assumed. Some restrictions will nevertheless be imposed later on the model to keep the result manageable, which effectively turns it into the standard Periodic Anderson Model (PAM) with on-site hybridization and on-site Coulomb interactions.

Similar to chapter 6 the f-orbital electrons are assumed in our model to be the localized electrons. These electrons are highly localized and can be used to replace the d-orbital electrons. This was originally proposed in Schrieffer and Wolff's paper for semiconductors. This difference makes the model particularly suitable for CuO$_2$ superconductor materials, in which conduction electrons are replaced by holes in the Oxygen while the localized f-orbital electrons are replaced by highly localized holes in the Copper atom. [55,97]

The unitary transformation as defined in the one dimensional case is $\tilde{H} = e^S H e^{-S} = H_0 + [S, H_V]/2 + [S, [S, H_V]]/3 + \ldots$, where $S$ is determined by the standard condition $H_V + [S, H_0] = 0$. We use the same equation as in the one dimensional case for $S$, in which

$$S = \sum_{i,j,\sigma}(A + Z f_{j,-\sigma}^1 f_{j,-\sigma}^i)(V_{i,j} f_{j,\sigma}^i c_{i,\sigma} - V_{i,j}^1 c_{i,\sigma}^1 f_{j,\sigma})$$

where $A = 1/(-2t + \mu_e + \epsilon_f)$ and $Z = 1/(-2t + \mu_e + \epsilon_f + U) - A$ as usual. Contrary to the one dimensional case, the Hamiltonian so far has no restriction on the interactions between electrons. In fact, the indices $i$ and $j$ in the kinetic term and the hybridization term can represent any site in the lattice, and the canonical
transformation would still work with the $S$ defined in Eq. (7.1). The first order transformation result ($n = 1$) with multiple site indices is therefore

$$H_1 = -2AV_{kj}c_{i\sigma}^\dagger c_{j\sigma} + 2AV_{ik}c_{i\sigma}^\dagger f_{j\sigma} - ZV_{ij}c_{k\sigma}^\dagger c_{j\sigma} f_{i\sigma} + ZV_{ij}c_{k\sigma}^\dagger f_{i\sigma} - 2ZV_{ij}c_{k\sigma}^\dagger f_{j\sigma} + ZV_{ik}f_{i\sigma} f_{k\sigma} f_{i\sigma}$$

(7.2)

As defined from the Hamiltonian, $V_{i,j}$ represents the hopping of an electron from the impurity site $i$ to the conduction site $j$. According to this, $V_{i,k}c_{j\sigma}^\dagger$ indicates the hopping of an electron from impurity site $i$ to impurity site $j$ via a conduction site $k$. In other words, all terms in the result show interaction of electrons between two sites via an intermediate step.

For higher order processes, the interaction using intermediate sites is even more obvious as the number of indices increases in the final result. The complexity of the result is also increased as a consequence. In some cases, this would give rise to long range interaction such as RKKY, double exchange or super exchange, but in other cases it might not have any significant impact to the physical properties when the intermediate steps do not involve any spin flips. We therefore need to find out in more detail the interaction with multiple sites. The first step towards this direction is to find the result with an arbitrary number of indices in the expression. It was relatively easy as shown above in Eq. (7.2) for the first order. When we went up to the third order, the result already became very involved as shown in Eq. (B.2) of Appendix B, which has taken up to four pages. This third order result, as well as the first order one, does not provide a great deal of interest to us as the form of the equations lacks some kind of pattern, the pattern that could lead us to predict the forms of higher order processes as in the one dimensional case. We therefore need to re-derive these results from a different angle and be able to extend them to higher order terms. Two approaches will be discussed in the following to understand these higher order processes. The first one is called non-crossing hopping which assumes particles never hop to the same site. This is true when the system is near the ground state and hopping is rare. The second case is called recursive on-site interaction which investigates the situation when electrons interact on a site by site basis. This provides a direct insight into the average characteristics of the model.

### 7.1.1 Non-crossing hopping limit

In this section, we start again from the basics of the commutation with the parameter $S$ to derive the full expression of the canonical transformation to any
order. All we need for this is just four commutation results. The first one is the commutation of $S$ with the hybridization term $H_V$.

$$[S, f_{i', \sigma}^\dagger c_{j', \sigma}^\sigma + c_{j', \sigma}^\dagger f_{i', \sigma}] = (A + Zn_{i, -\sigma}) (f_{i', \sigma}^\dagger f_{i', \sigma} \delta_{jj'} - c_{j', \sigma}^\dagger c_{j', \sigma} \delta_{ii''} + f_{i', \sigma}^\dagger c_{j', \sigma} \delta_{jj'} + Z(f_{i', \sigma}^\dagger c_{j', \sigma}^\sigma - c_{j', \sigma}^\dagger f_{i', \sigma})$$

The indices $i'$ and $i''$ are the same in the hybridization term of the Anderson model, but can be different in higher order terms. This is the first order result of the canonical transformation and therefore the same as in Eq. (7.2) but in a different form.

The second result we need is a commutation of a similar form with a sign change.

$$[S, f_{i', \sigma}^\dagger c_{j', \sigma}^\sigma - c_{j', \sigma}^\dagger f_{i', \sigma}] = (A + Zn_{i, -\sigma}) (-f_{i', \sigma}^\dagger f_{i', \sigma} \delta_{jj'} + c_{j', \sigma}^\dagger c_{j', \sigma} \delta_{ii''} + f_{i', \sigma}^\dagger c_{j', \sigma}^\sigma - c_{j', \sigma}^\dagger f_{i', \sigma})$$

This commutation appears in every odd order of the transformation, which will become very useful later in the formulation.

The third and the fourth results are

$$[S, A + Zn_{i, -\sigma}] = -Z(A + Zn_{i, -\sigma}) (f_{i', \sigma}^\dagger c_{j', \sigma} \delta_{ii''} + c_{j', \sigma}^\dagger f_{i', \sigma} \delta_{ii''})$$

$$[S, f_{i', \sigma}^\dagger f_{i', \sigma}^\dagger - c_{j', \sigma}^\dagger c_{j', \sigma}] = (A + Zn_{i, -\sigma}) (f_{i', \sigma}^\dagger f_{i', \sigma} \delta_{jj'} - c_{j', \sigma}^\dagger c_{j', \sigma} \delta_{ii''} - f_{i', \sigma}^\dagger f_{i', \sigma} \delta_{jj'} + c_{j', \sigma}^\dagger c_{j', \sigma} \delta_{ii''})$$

The third result produces a virtual process out of a f-electron number operator from the commutation, while the fourth result shows the commutation of a hopping term. The $i'$ and $i''$ indices in the fourth result are again identical in the first order result of the transformation, but can be different in higher order terms.

These four results look tedious and un-related in a first glance, but in fact they form a closed group. In other words, these are all the results that we need to determine terms of the transformation to any order we desire. It becomes more apparent if we associate the terms on the left hand side of each commutation as:

$$\alpha_{ij,j'} = f_{i, \sigma} c_{j, \sigma} + c_{j', \sigma}^\dagger f_{i, \sigma}$$

$$\beta_{ij} = A + Zn_{i, -\sigma}$$

and rewrite the four results as:

$$[S, \alpha_{ij,j'}] = \beta_{ij} (\gamma_{ij,j'} + \gamma_{i,j'}) + Z \omega_{ij,j'} \omega_{i,j'}$$

$$[S, \beta_{ij}^\sigma] = -Z \beta_{ij}^\sigma \alpha_{ij,j'} \delta_{ii''}$$

$$[S, \gamma_{ij,j'}] = -\beta_{ij}^\sigma (\omega_{ij,j'^'} + \omega_{i,j'^'}$$

$$[S, \omega_{ij,j'^'}] = \beta_{ij}^\sigma (\gamma_{ij,j'} - \gamma_{ij,j'^'}) + Z \omega_{ij,j'^'} \omega_{i,j'^'}$$

where the $\gamma$s and $\alpha$s are given by Eqs. (7.3).
One can easily read off from these expressions that commutation of $S$ with the hybridization term $H_V$ generates localized electron hopping $\gamma_{i''j''}$ as well as electron spin flipping $\omega^{-\sigma}\omega^\sigma$. We can carry this through to the second order and even the third order commutation processes, but as it stands the result would still be far from useful. The complication comes from the delta functions in the equation which link different processes together. The correlation of indices due to the delta function provides information about the physical path of the localized electrons moving around in the lattice and the locations of spin flipping. It however does not help particularly with the spin flipping process itself. To understand the spin interaction, we shall therefore assume an extreme case where all indices are different in this formulation, as opposed to another extreme case where all indices are identical in the next section. This assumption first of all sets the first term on the right hand side of Eq. (7.7) to zero, as the two $\gamma$ would yield the same result. In fact, even without this assumption, this term is still zero in the first order result due to the symmetry of the identical indices $i = i'$ and $j = j'$. The same assumption also simplifies the two additions in Eqs. (7.4) and (7.6), for the two operands of each addition being identical. Most of all, this assumption allows the pair operators $\alpha, \beta, \gamma$ and $\omega$ to move around in the equation without following the Pauli exclusion principle, as they would never be at the same site with the same indices. From another point of view, this assumption considers electrons hopping around in the lattice without running into each other, or without crossing. Since the indices are now irrelevant, we shall omit them from the formulation. For instance, a multiplication of two identical pair operators would mean two pair operators occurring at the same time but at different places. This should become clear from the context.

The second order term of the transformation can now be read off from the four equations Eqs. (7.4 - 7.7),

$$[S, [S, \alpha^\sigma]] = (-2Z\beta^{-\sigma}\gamma^\sigma - 4\beta^{-\sigma}\beta^{-\sigma} + 2Z^2\omega^\sigma\omega^{-\sigma})\alpha^\sigma$$ \hspace{1cm} (7.8)

The hybridization $\alpha^\sigma$ in this expression indicates that it is a virtual process, even though there is a spin flip term $\omega^\sigma\omega^{-\sigma}$ in the result. In fact every even order commutation with $S$ is a virtual process, with a multiplication of $\alpha^\sigma$ in the result.

The third order term is slightly more complicated, but in a similar form.

$$[S, [S, [S, \alpha^\sigma]]] = 12Z\alpha^\sigma\alpha^\sigma\beta^{-\sigma}\beta^{-\sigma} + 2Z^2\alpha^\sigma\alpha^\sigma\alpha^\sigma\beta^{-\sigma} - 8\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma5\alpha^\sigma}$$

$$-4Z\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma\gamma^\sigma + \omega^\sigma\omega^{-\sigma}(Z^2\alpha^\sigma\alpha^\sigma + Z^2\alpha^{-\sigma}\alpha^{-\sigma} + 2Z\beta^{-\sigma}\gamma^\sigma + 2Z\beta^\sigma\gamma^{-\sigma}$$

$$-2Z^2\beta^{-\sigma}\gamma^\sigma - 4Z\beta^{-\sigma}\beta^{-\sigma} + 2Z^2\alpha^\sigma\alpha^{-\sigma} + 2Z^2\omega^{-\sigma}\omega^\sigma)$$ \hspace{1cm} (7.9)

The spin flip term now reappears in the result without a virtual process attached to it, and the process now has gone up to four pair operators. Within the spin flip
term $\omega^\sigma \omega^{-\sigma}$, there are two more spin flips of the type $\omega$ and $\alpha$ happening at the same time, but at different sites. After expanding the $\alpha^\sigma \alpha^{-\sigma} + \omega^{-\sigma} \omega^\sigma$ term it is easy to see that the part corresponding to spin interaction does get cancelled out. This means the spin interaction is only of first order. Meanwhile, the impurity hopping $\gamma^\sigma$ has also reached the second order.

After working through the first three orders, we are now more confident in formulating the $n$th order result. To this end, we define the $n$th commutation of $S$ as

$$\Diamond^n S = [S, [S, [S, \ldots [S, X] \ldots]]]$$

(7.10)

which follows the addition and the multiplication rules of commutation as usual,

$$\Diamond^n S Y = \Diamond^n S Y + m \Diamond^{n-1} S Y + \cdots + X \Diamond^m Y, \Diamond^m (X + Y) = \Diamond^m X + \Diamond^m Y,$$

In terms of this new symbol, the $n$th order commutation of $S$ with $H_V$ can be written in a condensed form in which

$$\Diamond^{n+1} \alpha^\sigma = 2(\Diamond^n \beta^{-\sigma} \gamma^\sigma + C^1_1 \Diamond^{n-1} \beta^{-\sigma} \gamma^\sigma + \cdots + C^n_{n-1} \beta^{-\sigma} \Diamond^{n-1} \gamma^\sigma + \beta^{-\sigma} \Diamond^n \gamma^\sigma)$$

$$+ Z(\Diamond^n \omega^{-\sigma} \omega^\sigma + C^1_1 \Diamond^{n-1} \omega^{-\sigma} \omega^\sigma + \cdots + C^n_{n-1} \Diamond^{n-1} \omega^{-\sigma} \omega^\sigma + 2 \Diamond^n \omega^{-\sigma} \omega^\sigma + \omega^{-\sigma} \Diamond^n \omega^\sigma)$$

(7.11)

To solve this, we need to know the $n$th commutation of $\beta, \gamma$ and $\omega$. These three commutations are in fact very similar, due to the fact that their first commutation have the same form.

$$\Diamond \beta^{-\sigma} = -Z \alpha^\sigma \beta^{-\sigma}, \quad \Diamond \gamma^\sigma = -2 \alpha^\sigma \beta^{-\sigma}, \quad \Diamond \omega^\sigma = Z \alpha^{-\sigma} \omega^\sigma,$$

(7.12)

Solving one of them would basically solve the other two. Because of this, we will concentrate on the first equation with $\beta$. The first step is to express $\Diamond^{n+1} \beta, \Diamond^n \beta, \ldots, \Diamond \beta, \beta$ in terms of $\beta$ alone using linear algebra.

$$\Diamond^{n+1} \beta^{-\sigma} = -Z(C^n_1 \Diamond^n \beta^{-\sigma} \alpha^\sigma + C^1_1 \Diamond^{n-1} \beta^{-\sigma} \Diamond \alpha^\sigma + \cdots$$

$$+ C^n_{n-1} \Diamond^{n-1} \beta^{-\sigma} \Diamond^n \alpha^\sigma)$$

(7.13)

The equations for $\Diamond^n \beta^{-\sigma}, \ldots, \Diamond \beta^{-\sigma}$ can be built upon the same relation with different values of $n$ in Eq. (7.13). If we now define $B$ to be the column vector of $\Diamond^n \beta, \Diamond^n \beta, \ldots, \Diamond \beta$, and $C$ a constant column vector of zero except the last element as $\beta^{-\sigma}$,

$$B = [\Diamond^{n+1} \beta^{-\sigma}, \Diamond^n \beta^{-\sigma}, \ldots, \Diamond \beta^{-\sigma} \beta^{-\sigma}]^T, \quad C = [0, 0, \ldots, 0, \beta^{-\sigma}]^T,$$

(7.14)

where $T$ means the transpose of the vector. The set of linear equations to solve for $\Diamond^{i+1} \beta^{-\sigma}$ where $i = 0, 1, 2, \ldots, n$ can be written as

$$M \cdot B = C$$

(7.15)
in which \( M \) is the matrix containing the relationships between each element of \( B \), and is given by

\[
M = \begin{pmatrix}
1 & Z\alpha^\sigma & ZC_1^n\alpha^\sigma & ZC_2^n\alpha^\sigma & \cdots & Z\delta^n\alpha^\sigma \\
0 & 1 & Z\alpha^\sigma & ZC_1^{n-1}\alpha^\sigma & \cdots & Z\delta^n\alpha^\sigma \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 & Z\alpha^\sigma \\
0 & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}
\] (7.16)

Since the determinant of this matrix is 1, there exists an unique solution to the non-homogeneous matrix equation in Eq. (7.15). One can use Cramer’s rule to find the solution.

\[
\hat{\phi}^{i+1}\beta^{-\sigma} = \frac{|M_{n-i+1}|}{|M|} = |M_{n-i+1}| \quad \text{for } i = 0, 1, 2, \ldots, n,
\] (7.17)

where \( M_i \) is the same matrix as \( M \) except that the \( i^{\text{th}} \) column is replaced by the column vector \( C \). The determinant of \( M_i \) is not easy to find but what we do know about this determinant is that it is a function of \( \alpha^\sigma, \delta\alpha^\sigma, \cdots, \delta^n\alpha^\sigma \) only and it has a factor of \(-Z\beta^{-\sigma}\) in front. We shall therefore denote this determinant as

\[
-Z\beta^{-\sigma}f_i^\sigma(-Z) = |M_{n-i+1}|,
\] (7.18)

to emphasize its pre-factor where \( f \) is the determinant without the pre-factor. The function \( f \) is written explicitly in terms of \(-Z\) so that it can be re-used for \( \gamma \) and \( \omega \) later, when the only difference between the three equations in Eq. (7.12) is the constant multiplication factor in their commutations. In particular, the equation for \( \gamma \) can be obtained from the corresponding equation for \( \beta \) with the constant \(-Z\) replaced by \(-2\). This however, excludes the \( Z \) inside the first commutation of \( \alpha \) in Eq. (7.4) as this equation does not participate in the derivation. In fact \( \delta\alpha \) will be treated as one of the elementary variables until the final stage of the evaluation.

The same approach also applies to \( \omega \) whereby the constant \(-Z\) is replaced by \( Z \), the operator \( \alpha^\sigma \) is replaced by \( \alpha^{-\sigma} \) and the operator \( \beta^{-\sigma} \) is replaced by \( \omega^\sigma \).

In summary, we have obtained the equations we need.

\[
\begin{align*}
\hat{\phi}^i\beta^{-\sigma} &= -Z\beta^{-\sigma}f_i^\sigma \\
\hat{\phi}^i\gamma^\sigma &= -2\beta^{-\sigma}f_i^\sigma \\
\hat{\phi}^i\omega^\sigma &= Z\omega^\sigma f_i^{-\sigma} \\
\hat{\phi}^i\omega^{-\sigma} &= Z\omega^{-\sigma}f_i^\sigma
\end{align*}
\] (7.19)

After substituting these equations back to Eq. (7.11),

\[
\begin{align*}
\hat{\phi}^{n+1}\alpha^\sigma &= 2\beta^{-\sigma}(-Zf_n^\sigma(-Z)\gamma + C_1^n2Z\beta^{-\sigma}f_{n-1}^\sigma(-Z)f_1^\sigma(-2) + \cdots - 2\beta^{-\sigma}f_n^\sigma(-2)) \\
&+ Z^2\omega^\sigma\omega^{-\sigma}(f_n^{-\sigma}(Z) + C_1^nZf_{n-1}^{-\sigma}(Z)f_1^{-\sigma}(Z) + \cdots + f_n^{-\sigma}(Z)),
\end{align*}
\] (7.23)
This is still not the final solution for $\hat{n}^{n+1}\alpha^\sigma$ yet, but one can now write it only in terms of $\hat{n}\alpha, \hat{n}^{-1}\alpha, \ldots, \hat{\alpha}$ and $\alpha$, and theoretically solve the set of non-linear equations. We have included the results of the fourth and fifth orders in Appendix I for reference. As already known from the results of previous commutations, an additional operator is always added to the result every time the order is increased by one. For instance, the fourth order result contains terms of five hopping operators, whereas the fifth order result contains terms of six hopping operators\(^1\). The result in Eq. (7.23) therefore shows that there is always at least one spin flip term $\omega^\sigma\omega^{-\sigma}$ in the Hamiltonian. Multiple spin interaction is possible if there are more spin flip terms inside the bracket of the last term in Eq. (7.23). The third order result Eq. (7.9) for example contains only the single spin flip term as mentioned before, whereas the fifth order result Eq. (I.12) in Appendix I contains spin flip terms up to the third order. This is very similar to the scenario that we have encountered previously when we discussed the canonical perturbation theory in chapter 4. The only difference here is that there are no projection operators involved in the results. In other words, we allow for double electron creation and annihilation to happen from the first order result. The result here therefore encompasses a much wider range of different interactions between electrons and as a consequence contains a lot more different types of terms.

The result also contains electron movement due to either the hybridization process $\alpha^\sigma$, or the actual hopping process $\gamma^\sigma$. If nearest neighbour interaction is imposed onto the Hamiltonian, the fifth order result in Eq. (I.12) for example would show electrons interacting with each other up to five sites apart. It is possible to go to higher orders and find out more interactions between the electrons but so far it seems to be enough to get a glimpse of the different kinds of interactions involved in the non-crossing hopping. Because no approximation other than the non-crossing hopping is made in the derivation, the result contains all possible values of the model parameters if high enough orders are considered.

Although it is an advantage for us to know all the details of the interaction in the model, in many cases it is only the average that one can observe and is therefore interested in. For instance, one may like to find out the first order spin flip process of the whole model regardless of what other processes might have created. One way to find this is to select the appropriate processes and sum them all together. This would involve selecting the relevant processes out of a perturbative expansion which is most successfully done in the diagrammatic approach. It however requires the knowledge

\(^1\)A hopping operator contains a creation operator and an annihilation operator of the same kind which represents the process of hopping or spin flipping or both.
of the perturbative expansion to any order and the fulfilment of the summation convergence. Another possible way is to assume that particles would most likely to interact with other particles at the same site, and the first order effect of the average spin interaction must be dominated by this kind of interaction. This process is in fact exactly the same as diagonalizing all the real space indices and looking for the spin interacting term from the final Hamiltonian. This diagonalization procedure is discussed in detail in the next section, and for now it merely means setting all indices to the same value. Multiple spin flips in the Hamiltonian with the same real space index is identically zero, due to the Pauli exclusion principle, which does not contribute to the first order spin flips. Hopping terms in this case become number operators, but hybridization remains unchanged. This diagonalization procedure essentially turns the model into a one body problem. Because the diagonalization is applied only at the final result of a three dimensional model, the calculation still requires attention to the dimensionality of the model. We will see in the next section that the model can in fact be treated as a one body problem. This diagonalization procedure can also be considered as conduction electrons recursively interacting with the same impurity in the lattice and hence called recursive on-site interaction.

7.1.2 Recursive on-site interaction

The example towards the end of the last section considers the calculation of the first order spin interaction from the perspective of a repeated on-site interaction. This on-site interaction can however be derived from a completely different point of view with an identical result. Starting from an earlier result of the canonical transformation on a three dimensional model, the indices \( i, i' \) and \( j, j' \) in the results of the first order transformation in Eq. (7.2), or any order in general, can be any site in the lattice as no restrictions have yet been made on the physical proximity of the electrons during interaction. However, in the original Hamiltonian of Anderson Model it is normally assumed that the strongest interaction is among those electrons which are in close proximity. For instance, the hopping matrix is non-zero only for nearest neighbouring conduction electrons, and the strong on-site Coulomb interaction is effective only when the impurity electrons are close to each other. Even hybridization is normally an on-site interaction whereby conduction electrons and impurity electrons interact only if they are both at the same site. These assumptions are all based on the fact that the strongest interactions between electrons are of short distance in nature. We therefore assume that the major contribution of the result obtained above would be from terms that have interacting electrons at the same site. It is however not saying that all other terms such as
the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction, superexchange, double-exchange are negligible. These terms are nevertheless, considered as interaction in the nearest neighbourhood and, therefore, as first order correction. Terms with electrons in the next nearest neighbourhood would accordingly be considered as second order correction. Appendix C shows in more detail the impact of this short distance interaction on the Hamiltonian in terms of operators.

We will, therefore, assume that on-site interactions between electrons are much stronger than interactions between electrons of neighbouring sites, in which \( c_i^\dagger c_j \) is almost zero except when \( i = j \). The hybridization coefficient \( V_{i,j} \) will also have in a similar fashion which is what we normally see in the standard Anderson model anyway. The impact of this assumption on \( V_{i,j} \) can easily be modelled by a delta function peaked at \( i = j \) with a nominal magnitude \( V \). Another way of representing this dominant on-site interaction is to consider each term of the result in Eq. (7.2) as an element of a matrix, with two summing indices (such as \( i \) and \( j \)) being the indices for row and column. Terms with the third index (such as \( k \)) as intermediate steps between hopping will be summed within the element if possible, or treated as another independent index. The off-diagonal elements of this matrix therefore correspond to terms of cross-site interaction, while the diagonal elements correspond to terms of on-site interaction. The assumption that the on-site interaction is much stronger than that of the neighbouring sites, means that the sum of the diagonal elements are more substantial than the sum of all off-diagonal elements. It is therefore sufficient to the first order approximation to consider only the diagonal elements as the significant result of the transformation. We will in the following refer to this procedure as the diagonalization of the result, or index contraction as it imposes a strong restriction to the indices, and as a result reduces the number of summing indices to one.

Using the diagonalization procedure to simplify the first order result with a single summing index \( j \), we obtain for the first order result of the transformation,

\[
H_1 = \frac{1}{2} [\mathcal{S}, H_V] = \frac{\nu^2}{2} \sum_{j,\sigma} \left( 2(A + Zn_{j,-\sigma}^f)(n_{j,\sigma}^f - n_{j,\sigma}^c) + Z(2c_{j,\sigma}^\dagger c_{j,-\sigma}^{f\dagger} f_{j,\sigma}^f f_{j,-\sigma}^f - c_{j,\sigma}^\dagger c_{j,-\sigma}^{f\dagger} f_{j,\sigma}^f f_{j,-\sigma}^f - c_{j,\sigma}^\dagger c_{j,-\sigma}^{f\dagger} f_{j,\sigma} f_{j,-\sigma}^f) \right). \tag{7.24}
\]

This result is indeed identical to the result of the one dimensional case. The second and the third order results are also found to be the same as that of the one dimensional counterparts. The fact that the first three order results are the same as those of the one dimensional case is more than a coincidence. To understand
this, we first look at the difference between the three dimensional and the one dimensional model. In the one dimensional model, only one site index survives after the commutation with $S$ and therefore the result is always composed of only one index. In the three dimensional case, a new index is generated every time a commutation of $S$ takes place. Index contraction is therefore needed to reduce the complexity of the result and bring it down to a simpler form. This contraction is however applied only after the full result is obtained. The resemblance of the results between the one and the three dimensional cases therefore infers that applying the index contraction last as in the three dimensional case yields the same result as if the index contraction was applied after every commutation of $S$, as in the one dimensional case. In other words, the sequence of the commutation operation and the index contraction through Dirac delta functions, as proven in Appendix D, is entirely irrelevant. The third order result, as an example, can therefore be calculated from commuting the already simplified second order result with $S$, instead of commuting the second order result with $S$ before contracting the indices. This reduction of indices reduces the algebra drastically as the commutation operation, the most time consuming process, is now applied to an already simplified result. It can be verified that the fourth order and fifth order results after diagonalization are identical to the corresponding results in the one dimensional case in Eqs. (6.7) and (6.8).

According to this, we can use the result of the one dimensional case from chapter 6 for the three dimensional model. The physical significance of the result from chapter 6 will be discussed in the following, especially with respect to the spin interaction which has been extensively studied in the past.

### 7.2 Spin interaction

The term that is of most interest to us from the transformed Hamiltonian is undoubtedly the spin carrier interaction. We will first consider the symmetric case in which there are only two independent variables $V$ and $U$ in the Hamiltonian.

$$J_{\text{sym}} = \frac{8V^2}{U} \left( \frac{\sin(4V/U)}{4V/U} + \frac{\cos(4V/U) - 1}{16V^2/U^2} \right)$$

(7.25)

In the symmetric case, the spin-carrier coefficient $J$ has a simple form from which the second order, fourth order and other higher order terms in $V$ can be identified after Taylor expansion. The second order term is the same as the Schrieffer and Wolff's result for small $V$. As $V$ gets larger, the Schrieffer and Wolff's result is not
adequate because the third order term of the transformation becomes more significant, forcing the coefficient to diminish, to the order of \( V^4 \). This competing effect is significant only when \( V \) is outside the range specified in the Schrieffer and Wolff's paper, which means our result is still consistent with that of Schrieffer and Wolff's. The combined effect of the first and the third order transformation causes the coefficient \( J \) to have an asymmetric shape with respect to \( V \) as shown in figures 7.1 to 7.5. This behaviour of \( J \) is consistent with the results from experiments in which the Neel temperature for CePdSb and the Curie temperature for CeAg are plotted against pressure [27]. In the plots, the Curie temperature is assumed proportional to \( J^2 \) as Curie temperature is proportional to \( J_{RKKY} \), the spin exchange integral between two impurities, in the mean field theory [54] and \( J_{RKKY} \) is proportional to the square of spin carrier integral \( J \) between impurities and carriers. In the horizontal axis, the s-d and s-p conduction bands broaden under pressure which reduce the band gap between conduction band and localized f-orbitals, and therefore increase the hybridization between the two bands [84]. Figures 7.1 and 7.2 show how close the experimental Curie temperature of CePdSb are to the theoretical calculations. Figures 7.3, 7.4 and 7.5 show the theoretical fitting to the experimental results of three more different materials. Not all of them is as the best parameter fit but the shape of the theoretical estimation agrees well with that of the experimental results.

When \( V \) becomes larger, the third order effect will be modified by the fifth order result of the transformation which changes the behaviour of the interaction again. This effect however is not significant until \( 8V/U \) is larger than 2. The combined effect of the first, third and fifth order result is already enough to have a glimpse of an oscillation which is inherited from the trigonometric function of the result in Eq. (7.25) as explained in chapter 6. The same phenomenon also exists in most of the asymmetric cases.

This surprising result is however not unique to our investigation. E. Müller-Hartmann and A. Reischl [69] have observed a similar phenomenon when they derived the effective spin character from a three band model describing the low energy magnetic properties of undoped CuO\(_2\)-planes, using direct perturbation expansion up to the twelfth order with respect to the Copper-Oxygen hopping. The calculation started from a conventional three-band model with p-orbital Oxygen holes at four corners and a 3d-orbital Copper hole at the center of a unit cell. The Hamiltonian can be simplified as

\[
H = H_e + H_U + H_{pd}
\]  

(7.26)
7. The Three Dimensional Periodic Anderson Model

Figure 7.1: The curve fitting of the data points (x) from figure 2 of Cornelius, Gangopadhyay, Schilling [27] for CePdSb, using our spin-carrier coefficient \( J^2 \) with \( U = 8 \) and \( \mu_c = 0 \) (solid line).

where

\[
H_\epsilon = \sum_{k,\sigma} \epsilon_d d_{i,\sigma}^\dagger d_{i,\sigma} + \epsilon_p (p_{x,\sigma}^\dagger p_{x,\sigma} + p_{y,\sigma}^\dagger p_{y,\sigma}) \tag{7.27}
\]

\[
H_U = U \sum_k d_{i,\sigma}^\dagger d_{i,\sigma} + d_{j,\sigma}^\dagger d_{j,\sigma} \tag{7.28}
\]

\[
H_{pd} = -t_{pd} \sum_{i,\sigma} (d_{i,\sigma}^\dagger (p_{x,\sigma} + p_{y,\sigma} - p_{x,\sigma} - p_{y,\sigma}) + h.c.) \tag{7.29}
\]

The first term describes the energies of 3d and 2p orbitals involved, the second term describes the Coulomb repulsion on \( \text{Cu}^{+1} \) ions and the third term describes the hopping between 3d and neighbouring 2p orbitals. The 2p-orbital holes are equivalent to our conduction electrons while the 3d Copper holes are similar to our localized f-orbital electrons.

The perturbation theory given by Takahashi [81] provides a particular example and an explicit solution to an effective Hamiltonian of any order. Using this, Müller-Hartmann and Reischl [69] carried out a direct expansion of the Hamiltonian in
Figure 7.2: The curve fitting of the data points (x) from figure 3 of Cornelius, Gangopadhyay, Schilling [27] for CePdSb, using our spin-carrier coefficient $J^2$ with $U = 8$ and $\mu_c = 0$ (solid line).

Eq. (7.26) with respect to $t_{pd}$, in the undoped case where the ground states of the unperturbed Hamiltonian contain completely filled p-orbitals and a single d-hole on each Copper site. The effective spin Hamiltonian of the Copper holes was calculated up to the twelfth order in $t_{pd}$. The spin interaction coefficient carried a similar variation to our spin-carrier coefficient $J$ as shown in the asymmetric case in Appendix H (when we summed up to the equivalence of their twelve orders). The similarity is striking. A similar plot to Hartmann and Reischl’s result based on our equation is shown in figure 7.6 for comparison.

As pointed out in Hartmann and Reischl’s paper, the radius of convergence of the direct perturbation series considered was very small and hence their result was accurate only up to the point when the different order results started to diverge. Because of this, Müller-Hartmann and Reischl [69] used two other different methods, a Pade approximate and the Zhang and Rice [98] singlet model, to confirm the validity of the result in a restricted range of $V$. This limitation, however, does not exist in our transformation as all orders of the transformation have been taken into
consideration. In fact, one can see from figure 6.1 that the higher order terms of the transformation are normally zero for small $V$. They become significant only when $V$ is increased to a certain extent. This “extent” is different at different orders and increases when the order increases.

Many different canonical transformation have also been published in the past [100], [99], [16], [55] etc, for the spin-spin interaction of localized electrons (holes). For instance, Zhou and Zheng performed a similar canonical transformation, to find the RKKY interaction in the Periodic Anderson Model up to the order of $V^5$. They introduced an extended ‘Schrieffer Wolff’ transformation

$$\tilde{H} = e^{W_3}e^{W_1}He^{-W_1}e^{-W_3}.$$  \hspace{1cm} (7.30)

to obtain the effective Hamiltonian between interacting f-electrons. $W_1$ in the transformation is the first factor chosen to eliminate terms to the linear order of $V$, which is the counterpart of $S$ in our formulation, whereas $W_3$ is the second factor chosen to eliminate terms to the order of $V^3$. These factors are determined from...
Figure 7.4: The curve fitting of the data points (x) from figure 4 of Cornelius, Gangopadhyay, Schilling [27] for CeRh2, using our spin-carrier coefficient $J^2$ with $U = 8$ and $\mu_c = 0$ (solid line).

the following conditions:

$$H_1 + [W_1, H_0] = 0, \quad (7.31)$$
$$1/3[W_1, [W_1, H_1]] + [W_3, H_0] = 0 \quad (7.32)$$

The transformed Hamiltonian therefore becomes

$$H = H_0 + 1/2[W_1, H_1] + 1/8[W_1, [W_1, W_1, H_1]] + O(V^5) \quad (7.33)$$

where the second term contains terms of order $V^2$ while the third term of order $V^4$. The higher order terms in this formation are truncated and therefore neglected which restricted the range of parameters in the model for the rest of the calculation.

While our method does not provide the indirect spin interaction between localized f-electrons explicitly because of the index contraction in our formulation, our result does show the spin carrier interaction between the localized electrons and the carriers, which is consistent with Zhou and Zheng [99] results calculated to the second order of $V$. A similar transformation to Zhou and Zheng [99] was also used in
Zhou [100], but instead of using perturbative approximation, the author employed the Gutzwiller projection operators to project out the double localized f-electron occupancy. This method however still imposes a restricted range of $V$ and suffers from the neglect of higher order spin-carrier interaction, let alone the distortion introduced by the projection operators. In fact as pointed out by Zanen and Oles [97] which will be discussed later, the spin interaction between the localized electron and the carrier is one order of magnitude larger than the spin interaction between the localized electrons, and hence it is important not to underestimate the effect of the spin-carrier interactions.

Batista, Bonca and Gubernatis [16] have a similar approach to Zhou and Zheng [99], but instead of using a perturbative approach with Hartree-Fock approximation to obtain the RKKY interaction to the fourth order of $V$, they used another ‘Schrieffer Wolff’ transformation to eliminate terms to the second order of $V$. In this way, no perturbation and approximation are introduced and they obtained a final expression of the Hamiltonian with only spin interaction between the localized electrons. The assumption of a small $V$ is however still required for the higher order

Figure 7.5: The curve fitting of the data points (x) from figure 4 of Cornelius, Gangopadhyay, Schilling [27] for CeRh3, using our spin-carrier coefficient $J^2$ with $U = 8$ and $\mu_c = 0$ (solid line).
7.2. Spin interaction

Figure 7.6: The 3rd, 4th, 5th and 6th order of the spin-carrier interaction relative to the 2nd order in the canonical transformation, compared to the 6th, 8th, 10th and the 12th order spin-spin interaction relative to the 4th order in Hartmann, Reischl [69], with $U = 8$, $\epsilon_f = -3.6$ and $\mu_c = 0$.

terms of $V$ to be negligible.

Jan Zaanen and Andrzej Oles [97] generalized the canonical perturbation theory initially developed for the Hubbard Hamiltonian by Chao, Spalek and Oles [23], for the Anderson lattice model as already discussed in the canonical perturbation theory (CPT) in chapter 4. Using this canonical perturbation theory, Zaanen and Oles obtained the spin-spin (superexchange), carrier-carrier and spin-carrier interaction to the fourth order of the perturbation and showed that the spin-spin interaction is always considerably smaller than the second order spin-carrier interactions. The second order spin-carrier interaction they obtained is identical to our result when only on-site interaction is considered. This result is also confirmed by other papers such as Kolley and Kolley. [55]

The fourth order spin-carrier interaction of the CPT from Zaanen and Oles [97] is however different to what we obtained. The difference between the two results are obviously due to the restriction imposed by the subspace projection in the CPT. To understand this difference, we would need to go back to the definition of
"orders" in both formulations. In Zaanen and Oles' paper [97], they defined the zero order as the ground state in which there is always one hole located on each Cu site throughout the whole lattice. The second order of the canonical perturbation theory then considers all possible hopping of holes to the excited states which have only one site either doubly occupied or completely empty. Similarly the fourth order process considers excited states with two sites available for double occupancy or zero occupancy, and the perturbation is calculated based upon all possible single electron hopping to these excited states.

In our formulation, the order of result is defined by numbers of commutations of $S$ with the perturbed Hamiltonian $H_V$, as defined in Eq. (6.9). For instance, the first order term is the result of $S$ commuted with $H_V$. This term is proportional to $V^2$ and hence equivalent to the second order result of Zaanen and Ole's. Physically, this "order" can also be considered as the number of times electrons hybridizing between localized and conduction bands. In a second order process, an electron can hybridize twice to return to its original state. This consists of either a free electron hybridizing to a localized electron band and then hopping back, or a localized electron hybridizing to the conduction band and returning. These processes are identical to the corresponding electron hopping in the second order process of the canonical perturbation theory in Zaanen and Oles [97] and therefore the two results are identical.

In the fourth order process, our result considers a broader range of interaction to that of Zaanen and Oles [97]. The hybridizing electrons in Eq. (6.7) considered in the commutation of $S$ with the third order result in Eq. (6.5), include those electrons which start from four different states originated from the third order process. These four states in fact originate from the results of the first order process which include spin-carrier interactions and pair electron interactions. In other words, the fourth order result considers hybridizations not only from states which are the result of single electron hopping, but also states that are the result: of electron pair creation and annihilation, which are forbidden in Zaanen and Oles' approach [97]. Figure 7.7 shows a particular example in which two conduction electrons are hybridized to two localized sites with double occupancy during the second order process (on the left of the figure). The third order process (in the middle of the figure) then turns one of the doubly occupied localized electrons back to a singly occupied state, while the fourth order process (on the right of the figure) turns the other one to a singly occupied state. Meanwhile, the spin of the electrons in the process can be different, creating a spin coupling between the two types of electrons and therefore producing a spin interaction. This process is however not included in Zaanen and Oles's canonical
perturbation theory as pair electron creation or destruction is prohibited from the projection operation.

\[
\begin{align*}
&\text{I) Pair Hybridization} & \text{II) Single hybridization of different spin} & \text{III) Single hybridization of the same spin} \\
&\text{Process I on the left describes the creation of two localized electrons from two free electrons. Process II in the middle hybridizes one of the localized electrons of opposite spin to free electrons. Process III on the right hybridizes the other localized electron of the same spin to free electron.}
\end{align*}
\]

Another way of using projection operators to simplify the results of a canonical transformation is employed by E. Kolley, W. Killey and R. Tictz [55] who used projections to show an effective model Hamiltonian for describing the interplay between magnetism and superconductivity in the Cu-spin and O-hole subsystems using the Periodic Anderson Model. They carried out the Schrieffer and Wolff transformation of the Anderson lattice model up to the fourth order before projecting the result onto the subspace of singly occupied Cu sites. This procedure is not invertible, as emphasized in their paper that the projection must be carried out only at the end of the calculation. In fact their projection would project out the spin-carrier interaction in the Hamiltonian that we are trying to obtain for comparison and hence the result would have been rendered useless to us. Fortunately, the full result before projection was also published which allowed us to compare our result as a special case of theirs. The two results were indeed found to be identical. The full details of the comparison are collected in Appendix F. The consistency is completely to our expectation since the results were obtained from the same transformation, with one specializing in the spin-spin interaction between Cu-holes while the other concentrating on the spin-carrier interaction. It is also not surprising that their result, like ours, does not yield the same set of coefficients as those from Zaanen and Oles [97] as electron pair creation and annihilation is allowed in their formulation, except in the final projection stage.

All of the references we mentioned above however have one feature in common. They never mentioned the possibility of oscillation that could be derived from their results. This oscillation is an interesting phenomenon as described early in this chapter and in chapter 6. First of all, the frequency of this oscillation increases as the
Hamiltonian approaches the poles of the coefficients $A$ and $Z$, which suggests that the oscillation must be related to the transformation, or the ways the parameters are chosen in the transformation. For instance, if the coefficients $A$ and $Z$ were pole-free, the singularities in the results of the transformation would never have existed. The poles in the coefficients however depend on the conditions we chose to impose upon the transformation. In chapter 6, the parameter $S$ in the transformation was chosen in such a way that the first order hybridization term in the Hamiltonian was cancelled. Choosing a condition other than that would change the expression of $S$ and inevitably the results of the transformation. In the next two sections, we will explore the possibilities of imposing a different condition other than the one chosen by Schrieffer and Wolff in chapter 6, to eliminate not just the first order but all orders of hybridization in the Hamiltonian. It turned out that the poles in the coefficients $A$ and $Z$ disappear, and the oscillation is totally removed from the Hamiltonian.

### 7.3 Elimination of higher order hybridization terms

So far, the canonical transformation is performed based on the approach of Schrieffer and Wolff, eliminating the first order hybridization term in the Hamiltonian. As already shown in the result of the second order commutation of $S$ with $H_V$ in Eq. (6.3), hybridization terms reappear from the transformation, despite in higher orders of $V$. When $V$ is small, the higher order terms will be negligible and hence the hybridization is approximately zero. However for other values of $V$, such as those in the mixed valence region, this approximation is no longer valid and a better method is needed to completely eliminate the hybridization process.

The approach of Schrieffer and Wolff in the last section does not remove all the hybridization terms in the Hamiltonian, simply because the values of $A$ and $Z$ were chosen 'only' to remove the perturb Hamiltonian $H_V$, not the sum of the hybridization terms from the transformation. If this sum of the hybridization terms is known, it will be possible to choose $A$ and $Z$ in such a way that this sum will be identically zero. In other words, if one can express the transformed Hamiltonian up to infinite order and leave the parameters $A$ and $Z$ free, one can solve for $A$ and $Z$ by setting this sum to zero. This is precisely the motivation for the following calculation.

To evaluate the transformed Hamiltonian without fixing $A$ and $Z$, we would
7.3. Elimination of higher order hybridization terms

have to calculate all the terms from the canonical expansion:

\[ e^S H e^{-S} = H_0 + [S, H_0] + \frac{1}{2!} [S, [S, H_0]] + \ldots \]

\[ + H_V + [S, H_V] + \frac{1}{2!} [S, [S, H_V]] + \ldots \] (7.34)

without any reduction of terms from using the condition in Eq. (5.95). Since the
calculations of [S, H_V], [S, [S, H_V]], \ldots in the last section were done based upon
the most generic expression of A and Z, we can re-use the corresponding results without
imposing any condition on A and Z. All other terms of the transformation such as
[S, H_0], [S, [S, H_0]], \ldots, would however need to be calculated.

We begin with the first commutation,

\[ [S, H_0] = \sum_j V(A' + Z'n^j_{-\sigma})(f^j_{\sigma} c_{\sigma} + c^j_{\sigma} f_{\sigma}) \] (7.35)

in which

\[ A' = (2t - \mu_e - \epsilon_f)A, \] (7.36)

\[ Z' = (2t - \mu_e - \epsilon_f)Z - U A - UZ, \] (7.37)

This result is very similar to that of Eq. (6.1) and certainly falls into the pattern
of the even order commutation in Eq. (6.10) of chapter 6 which, as already calculated,
gives the next order commutation as:

\[ [S, [S, H_0]] = \sum_{i,\sigma} \left(J'_0(c^i_{\sigma} c_{i,-\sigma} f^i_{i,-\sigma} f_{i,\sigma} - n^i_{i,\sigma} n^i_{i,-\sigma}) + G'_0(n^f_{i,\sigma} - n^f_{i,-\sigma}) \right) \]

\[ + P'_0(c^i_{\sigma} c^i_{-\sigma} f_{i,\sigma} f_{i,-\sigma} + f^i_{i,\sigma} f^i_{i,-\sigma} c_{i,\sigma} c_{i,-\sigma}) + I'_0 n^f_{i,-\sigma} n^f_{i,\sigma} \] (7.38)

in which

\[ J'_0 = 2V^2(A'Z + Z'A + Z'Z); \] (7.39)

\[ P'_0 = -V^2(A'Z - AZ'); \] (7.40)

\[ G'_0 = 2V^2 AA'; \] (7.41)

\[ I'_0 = 2V^2(A'Z + Z'A + Z'Z); \] (7.42)

A prime is added to each coefficient to differentiate it from that of chapter 6.
This result is indeed very similar to the result of the first order transformation from
chapter 6, albeit different values of A, Z, J_0, P_0, G_0 and I_0. The commutation of H_0
with S in the higher order result is also the same as those of the one dimensional
case, if we consider a different set of coefficients J, P, G, I, K, M and A, Z. The
underlying algebra is identical provided the coefficients are generic. This allows us
to re-use the proof for the one dimensional model with different initial conditions and different values of $A$ and $Z$. To be consistent with the previous results, Eq. (7.38) will be considered as the first order result (i.e. $n = 0$) and $[\mathcal{S}, H_0]$ as part of the unperturbed Hamiltonian. Accordingly, the $n$th order commutation of $[\mathcal{S}, H_0]$ is

$$[\mathcal{S}, [\mathcal{S}, H_0]]_n = \sum_{m,i,\sigma} (J_m (c_{i,\sigma}^\dagger c_{i,-\sigma} f_{i,-\sigma}^\dagger f_{i,\sigma} - n_{i,\sigma} c_{i,-\sigma})$$

$$+ P_m (c_{i,\sigma}^\dagger c_{i,-\sigma} f_{i,-\sigma}^\dagger f_{i,\sigma} + f_{i,\sigma}^\dagger f_{i,-\sigma} c_{i,\sigma} c_{i,-\sigma} + G_m (n_{i,\sigma} c_{i,-\sigma} - n_{i,-\sigma})$$

$$+ I_m n_{i,-\sigma} n_{i,\sigma} + M_m n_{i,-\sigma} n_{i,\sigma} (n_{i,\sigma} - n_{i,-\sigma} + K_m n_{i,-\sigma} n_{i,\sigma}) ),$$

(7.43)

where $n = m + 1$ for odd values of $n$ and non-negative values of $m$, as opposed to Eq. (6.10) in the Schrieffer and Wolff’s approach. The values of the coefficients are given by Eqs. (6.25,6.28,6.29,6.30,6.31,6.32), with the initial conditions stated in Eqs. (7.39-7.42). For even values of $n$,

$$[\mathcal{S}, [\mathcal{S}, H_0]]_n = \sum_{m,i,\sigma} (R_m + S_m n_{i,-\sigma} + T_m n_{i,-\sigma}^c + Q_m n_{i,-\sigma}^c)$$

$$\cdot (c_{i,\sigma}^\dagger f_{i,\sigma} + f_{i,\sigma}^\dagger c_{i,\sigma}) ,$$

(7.44)

where $n = 2m + 2$ and

$$R_m = -2VC_m^a A ,$$

(7.45)

$$S_m = -2V[(J_m + P_m)A + I_m (A + Z) + (V_m + G_m)Z] ,$$

(7.46)

$$T_m = 2V[(J_m + P_m + K_m)A + P_m Z] ,$$

(7.47)

$$Q_m = 2V[(J_m + K_m - P_m)Z - M_m (A + Z)] .$$

(7.48)

These coefficients are calculated with the new set of initial conditions given by Eqs. (7.39-7.42).

The total Hamiltonian is therefore

$$\tilde{H} = H_0 + H_V + [\mathcal{S}, H_0] + \sum_{n=1} \frac{1}{(n + 1)!} [\mathcal{S}, [\mathcal{S}, H_0]]_n + \sum_{n=1} \frac{1}{(n)!} [\mathcal{S}, H_V]_n$$

(7.49)

The first three terms can be expanded using Eqs. (4.1, 4.2, 7.38) while the last two terms have to be be calculated using the trigonometric functions we used in chapter 6. It is however easier to find out the coefficient of each interacting term in the Hamiltonian if it is written in the format of Eq. (6.45) using the expressions for
7.3. Elimination of higher order hybridization terms

$H_{even}$ and $H_{odd}$ defined below.

$$H_{odd} = \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} \{S, H_V\}_{2m+1} + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} \{S, [S, H_0]\}_{2m+1}$$  \hspace{1cm} (7.50)

$$H_{even} = H_V + \{S, H_0\} + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} \{S, H_V\}_{2m+2}$$

$$+ \sum_{m=0}^{\infty} \frac{1}{(2m+3)!} \{S, [S, H_0]\}_{2m+2}$$  \hspace{1cm} (7.51)

These new even and odd Hamiltonians can be summed to all orders, similar to the ones in chapter 6, to yield Eqs. (6.48) and (6.55) with different values of $J, P, G, M, I, K$ for the odd order and $R, S, T, Q$ for the even order. For instance, the coefficient $R$ in $H_{even}$ contains terms from $H_V, [S, H_0], R'_m$ and $R_m$.

$$R = V + A'V + \sum_{m=0}^{\infty} \frac{(-4V^2A)^m}{(2m+2)!} (-2VAG_0) + \sum_{m=0}^{\infty} \frac{(-4V^2A)^m}{(2m+3)!} (-2VAG'_0)$$  \hspace{1cm} (7.52)

Using the Taylor expansion of sine and cosine functions with the initial condition given by Eqs. (7.39-7.42), one can rewrite it as

$$R = V + A'V + (\cos(2VA) - 1)V + (\sin(2VA)/2VA - 1)VA'$$  \hspace{1cm} (7.53)

Now setting this to zero, instead of just the first order term $V + A'V$ as in the case of Schrieffer and Wolff, one can solve for $A$ as

$$A = \frac{1}{2V} \tan^{-1} \left( \frac{-2V}{2t - \mu_c - \epsilon_f} \right)$$  \hspace{1cm} (7.54)

If $V$ is small relative to $t, \mu_c$ and $\epsilon_f$, the arctan function can be approximated by a linear function and $A$ becomes

$$A = \frac{-1}{2t - \mu_c - \epsilon_f}$$  \hspace{1cm} (7.55)

This is just what we defined as $A$ in the last section and also the corresponding coefficients in the Schrieffer and Wolff transformation. Therefore, Schrieffer and Wolff transformation can be considered as a special case of this transformation when $V$ is small.

On the contrary when $V$ is large or $\epsilon_f$ is close to zero, this new result of $A$ makes a substantial difference to all the coefficients determined from the transformation. Thanks to the arctan function in the result of $A$ in Eq. (7.54), this new value of $A$ will never go infinite. In other words, a singularity is removed from the transformation as the pole in $A$ at $\epsilon_f = 0$ in the case of Schrieffer and Wolff (and our result from chapter 6) is replaced by an arctan function which maps the entire real
axis onto an interval of $\pm \pi/2$ modular $\pi$. The singular point is therefore mapped to $\pm \pi/2, \pm 3\pi/2, \pm 5\pi/2, \ldots$ depending on the values of the function immediately before and after the singular point for continuity. This also shows that the oscillatory nature of the results in chapter 6 is related to the hybridization term of the Hamiltonian. In the case of Schrieffer and Wolff when only the first order hybridization is removed, singular points are introduced together with hybridization in the higher order of the transformation. Effectively speaking, the transformation weighs and re-shapes the overall hybridization in such a way that the coefficients are highly suppressed in the region when $V/U$ is small, but relaxed in all other region of the parameter space, especially at the region near the singularities. The higher order results of the transformation impose even more suppression on the coefficients at the singular points, where the first order singularities are now turned into higher order oscillations. Only when the final sum of the hybridization terms are set to zero would the singularities be eventually removed, which indicates that the singularity is an artefact of the transformation. Looking from another point of view, the oscillation is due to a small residue remaining in the imaginary axis of the Hilbert space after the model is 'rotated' by the transformation to remove the first order hybridization.

Before we proceed to the next step, there is in fact another pole at $U = \epsilon_f$ in the result introduced by the coefficient $Z$ in Eq. (6.1) when $t$ and $\mu_c$ are set to zero. This pole, as suggested from the expression in Eq. (6.1), is related to the hybridization of a slightly more complex form $n'_i \sigma (f_i' \sigma c_i \sigma + c_i' \sigma f_i \sigma)$. The corresponding coefficient $S$ in the final summation would therefore need to be set to zero to eliminate this form of hybridization and the poles associated with it. This condition eventually leads to the determination of $Z$, as in the result of Schrieffer and Wolff. In the case when $A$ and $Z$ are free, the coefficient $S$ can be derived in the most general form with all orders taken into consideration, in terms of $[S, H_3], S_m'$, and $S_m$, similar to the coefficient $R$.

\[
S = VZ' + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} S_m' + \sum_{m=0}^{\infty} \frac{1}{(2m+3)!} S'_m
\]  

(7.56)

After substitution, one gets

\[
S = VZ' + \phi(\sqrt{8V^2(\alpha^2 + \beta^2)})4\alpha J_0 V + \phi(\sqrt{2V^2(\alpha^2 + \beta^2)}) \cdot \left( \frac{\beta^2}{\alpha^2 - \beta^2} J_0 - \beta \frac{\beta^2}{\alpha^2 - \beta^2} P_0 \right) 2V - \phi(2V\beta)2V\beta G'_0 + \psi(\sqrt{8V^2(\alpha^2 + \beta^2)}) \cdot 8\alpha(\alpha - \beta)V^3 + \psi(\sqrt{2V^2(\alpha^2 + \beta^2)})2\beta(\alpha + \beta)V^3 - \psi(2V\beta)4\beta^2 V^3
\]  

(7.57)

in which $\phi(x) = (\sin(x) - x)/x^3$ and $\psi(x) = (\cos(x) - 1)/x^2$ and $\alpha = A + Z, \beta = A$ as defined earlier in this chapter. Setting $S$ to zero requires solving a trigonometric
7.3. Elimination of higher order hybridization terms

The equation which can be quite complex, considering the equation is highly non-linear. However, we can get around this by choosing a large value of \( Z \) relative to \( A \), using the fact that the equation is periodic and crossing the \( Z \)-axis at large \( Z \), as shown later. In fact the vertical offset of the periodic function is always zero, and hence the zero crossing points are evenly distributed along the \( Z \)-axis. If we assume \( Z \) is much larger than \( A \), we can simplify the equation by using

\[
\sqrt{\alpha^2 + \beta^2} \approx A + Z = \alpha \quad \beta / \alpha \approx 0,
\]

and get

\[
S = \frac{2t - \mu_c - \epsilon_f - U}{2\sqrt{2}} \sin(2\sqrt{2}\alpha V) + V \cos(2\sqrt{2}\alpha V) \\
- \left( \frac{2t - \mu_c - \epsilon_f}{2} \right) \sin(2\beta V) - V \cos(2\beta V).
\]

Setting this equation to zero and solving for \( \sin(2\sqrt{2}\alpha V) \) and \( \cos(2\sqrt{2}\alpha V) \) is simple. The solutions for \( S = 0 \) are

\[
\sin(2\sqrt{2}\alpha V) = \frac{\kappa \lambda + V\sqrt{\lambda^2 + V^2 - \kappa^2}}{\lambda^2 + V^2},
\]

\[
\cos(2\sqrt{2}\alpha V) = \frac{\kappa V - \lambda\sqrt{\lambda^2 + V^2 - \kappa^2}}{V^2 + \lambda^2},
\]

where

\[
\kappa = \frac{2t - \mu_c - \epsilon_f}{2} \sin(2V A) + V \cos(2V A),
\]

\[
\lambda = \frac{2t - \mu_c - \epsilon_f - U}{2\sqrt{2}}.
\]

The signs in front of the square roots in the solutions are determined by the initial condition when \( V \) is very close to zero. The validity of this approach hinges on the assumption that these solutions of \( Z \) are real, and hence the equation \( S \) does cross the \( Z \)-axis at some real value of \( Z \). In other words, one needs

\[
\lambda^2 + V^2 - \kappa^2 > 0
\]

in the solution. After substitution, this inequality becomes

\[
\frac{(2t - \mu_c - \epsilon_f - U)^2}{8} + V^2 - \frac{(2t - \mu_c - \epsilon_f)}{2} \sin 2V \beta + V \cos 2V \beta^2 > 0
\]

Using the solution of \( A \) from Eq. (7.54) and noting that \( \beta = A \), the last bracket becomes identically zero and therefore the inequality is always true. That means the solution of \( Z \) from Eqs. (7.60) and (7.61) are always real and the assumption
we made earlier in the paragraph is verified. The assumption is however needed only for solution finding and ease of plotting. The Eq. (7.57) can in fact be solved by numerical methods for other values of \( Z \) if necessary.

Using the solution of \( A \) from Eq. (7.54), \( \kappa \) is identically zero as pointed out before and hence the sine and cosine functions in Eqs. (7.60) and (7.61) are always less than or equal to one. The value of \( Z \) in the argument of the sine and cosine function is therefore real. Also, the expressions on the right hand side of Eqs. (7.60) and (7.61) are analytical for all values of \( U \) and \( V \), except possibly when \( U = 0 \) and \( V = 0 \) which would however turn the system into a free electron model and is therefore not interacting. As a result, \( Z \) is pole-free and will never become infinite. Together with a similar condition of \( A \), it is enough to see from the result that all coefficients of the Hamiltonian are analytical and free of singularities.

The symmetric case, when \( \epsilon_f = U/2 \), normally offers a much simplified picture and clearer solution to the Hamiltonian, as opposed to the more generic asymmetric case. It is however not the case here when only coefficients \( R \) and \( S \) are set to zero. In the Schrieffer and Wolff transformation, the symmetric case yields \( A = -Z \). Even in the next section when we consider all even coefficients \( R, S, T \) and \( Q \) to be zero, the symmetric case also reduces to \( A = -Z \). It is however not possible in this case to get \( A = -Z \) from setting \( S = 0 \) in the symmetric condition. The reason for this is believed to be the asymmetrical relationship between the conduction and the localized electrons we placed on the hybridization during the transformation. In the Schrieffer and Wolff case when only the first order hybridization is removed, no restriction is placed on any higher order hybridization. As far as the first order result is concerned, all hybridizations are removed. In the next section when all hybridizations are actually removed to all orders, no bias is set among the hybridization terms. Only in this case when two coefficients \( R \) and \( S \) out of four \( R, S, T \) and \( Q \) are set to zero, the hybridization of the form \( f^\dagger_{j\sigma} c_{j\sigma} + c^\dagger_{j\sigma} f_{j\sigma} \) and \( n_{j\sigma}(f^\dagger_{j\sigma} c_{j\sigma} + c^\dagger_{j\sigma} f_{j\sigma}) \) are singled out and removed, but other hybridization terms remain. This places a preference among some of the hybridization terms and hence causes an imbalanced condition in the symmetric case. In spite of this, we still consider the symmetric case in the following, for comparing with the result from chapter 6 and the result of Schrieffer and Wolff.

Before solving for \( Z \), it is possible to evaluate the coefficient \( J \) of the spin-carrier
7.3. Elimination of higher order hybridization terms

Figure 7.8: The spin-carrier interaction $J/V$ vs. $U/8V$, with $t = 0, \mu_c = 0$ and $V = 1$ in the symmetric case when $R = S = 0$ (solid line) and $R = S = T = Q = 0$ (long dashed line). The insert shows the same function vs. $8V/U$, when $U = 6$. Grey short dashed is the Schrieffer and Wolff’s result and grey long dashed is our 1D result.

Using the initial condition given by Eqs. (7.39-7.42) with the approximation of $\alpha$ and $\beta$ in Eq. (7.58), $J_{sym}^{(2)}$ becomes

$$J_{sym}^{(2)} = \frac{2t - \mu_c - \epsilon_f - U}{4} \left(1 - \frac{\kappa V + \lambda \sqrt{1 + (V^2 - \kappa^2)/\lambda^2}}{V^2 + \lambda^2}\right)$$

$$+ \frac{V(\kappa \lambda - V \lambda \sqrt{1 + (V^2 - \kappa^2)\lambda^2})}{\sqrt{2}(\lambda^2 + V^2)}$$

(7.67)

If the sign of the square root is chosen correctly from Eq. (7.61), one would get $J \to 0$ when $V \to 0$, which agrees with the physics of the model. Figure 7.8 shows
this coefficient (solid line) in the symmetric case when $V = 1$. $U = 6$ is chosen for the plot in the insert.

The difference between this value of $J^{(2)}/V$ and the value of $J/V$ from chapter 6 is clear. The sinusoidal behaviour is absent in the system when the coefficients $R$ and $S$ are set to zero. The new $J^{(2)}/V$ starts from a fixed value $1/\sqrt{2}$ at $U = 0$ (solid line) and decays in the same fashion as the maximum of the old $J/V$ (grey long dashed line) from chapter 6. The result of Schrieffer and Wolff is also shown for comparison in grey short dashed line. It is clear from the figure that the singularity is replaced by a fixed constant at $U = 0$, but the new $J^{(2)}/V$ does not coincide with the upper bound of the ripple in the old $J/V$. It would have been the case if $A = -Z$, as will be seen in the next section, but in this case the imbalanced condition imposed upon the hybridization terms causes the coefficient $J^{(2)}/V$ to drop slightly. It is more clearly shown from the insert of the figure when the coefficient $J^{(2)}/V$ is plotted against $V$.

Now if we convert the electron operators to spin operators as follows

$$S^+_c = c^+_c c_1, S^+_f = f^+_f f_1, S^-_c = c^+_c c_1, S^-_f = f^+_f f_1$$

$$S^+_c = \frac{1}{2} (c^+_c c_1 - c^+_1 c_c), S^-_f = \frac{1}{2} (f^+_f f_1 - f^+_1 f_f)$$

It is easy to see that

$$\sum_\sigma J(c^+_\sigma c_\sigma f^+_\sigma f_\sigma - n^c_\sigma n^f_\sigma) = 2J(S^c \cdot S^f - \frac{1}{4} n^c n^f)$$

which indicates spin interaction between localized and conduction electrons are antiferromagnetic, as $J$ is always positive from our calculation.

Our approach for $J$ can also be applied to the other coefficients of the Hamiltonian, to get the complete picture of the system. The first one is the pair hybridization.

$$P^{(2)} = -V \frac{\sin \sqrt{2}V\alpha}{\sqrt{2}} \quad (7.68)$$

This new $P^{(2)}/V$ (solid line) coefficient is shown in figure 7.9 which increases along the lower bound of the old $P/V$ (grey long dashed line) function from chapter 6. The result of Schrieffer and Wolff is also shown in grey dotted line for comparison. This new $P^{(2)}/V$ starts from a constant value $-0.5$ at $U = 0$ and increases as $U$ increases, overlaying the lower bound of the old $P/V$. The insert of figure 7.9 shows the same function of $P^{(2)}/V$ against $V$. It shows clearly that the new $P^{(2)}/V$ sites at the lower bound of the old $P/V$ function.

This pair term associated with the coefficient $P$ carries a very similar appearance to the Cooper Pair in the subject of superconductivity. It describes the creation
of a localized spin pair, following the annihilation of a conduction spin pair, and vice versa. Our pair term is however formulated in real space and localized at the same site of the lattice, which essentially is different to the Cooper Pair. It nevertheless does indicate the relative importance of pair creation and annihilation in the Hamiltonian which should not be overlooked, as we explained early in the chapter.

\[
G^{(2)} = \frac{2t - \mu_c - \epsilon_f}{2} (1 - \cos 2V\beta) + V \sin 2V\beta
\]  \hspace{1cm} (7.69)

The coefficient \(G\) has a very similar feature as the coefficient \(P\) described before. The new coefficient \(G^{(2)}/V\) sits at the lower bound of the old \(G/V\) function from chapter 6. It starts with a constant \(-1\) at \(U = 0\) and slowly decays to zero. Also shown in the figure is the result of Schrieffer and Wolff in grey short dashed line. The insert of the figure shows the same function against \(V\) which indicates how stable the function is over \(V\).
The $G$ coefficient as defined in Eq. (7.43) reflects the difference of the impurity energy level from the conduction band, generated from the transformation. When these new terms are renormalized to the original Hamiltonian, it shows the true energy level of the localized and the conduction electrons. Since the conduction energy level is assumed zero in the figure, the negative value of $G$ defines the new conduction energy level. With the localized energy level, the renormalization procedure essentially adds an extra factor $\epsilon_f$ to $G$ as shown in figure 7.11. In the symmetric case when $\epsilon_f = -U/2$, this extra factor of $\epsilon_f$ causes $G/V$ to decrease linearly with $U$ at large $U$, due to the dominating $\epsilon_f$ from the unperturbed Hamiltonian which drives the system into a local moment regime.

\[
I^{(2)} = \frac{(2t - \mu_c - \epsilon_f - U)}{4} \left(1 - \cos 2\sqrt{2}aV\right)
+ \frac{V}{\sqrt{2}} \sin 2\sqrt{2}aV - G^{(2)}
\]  

(7.70)
7.3. Elimination of higher order hybridization terms

The coefficient $I^{(2)}/V$ is plotted in figure 7.12 with solid line which shows the Coulomb interaction between the localized electrons generated by the transformation. It starts off with a constant $\approx 1.7$ at $U = 0$ and decays to zero. Similar to the $J$ coefficient, the new $I^{(2)}/V$ function does not sit on the upper bound of the old $I/V$ (grey long dashed line) from chapter 6, for the same reason as in the $J$ coefficient. They are however very close and in fact indistinguishable at large $U$. The insert of the figure shows the same function against $8V/U$ when $U = 0$. It indicates how stable the function is over $V$ and how close it is to the upper bound of the ripple from the old $I/V$ function.

This new value of the Coulomb interaction should in fact be combined with the one from the unperturbed Hamiltonian, to see the overall effect on the Hamiltonian. Since the Coulomb interaction in the unperturbed Hamiltonian is only $U$, the new renormalized Coulomb interaction would simply be $I^{(2)} + U$ as shown in figure 7.13. Similar to the renormalized $G$, the Coulomb interaction is dominated by the linear...
term from the unperturbed Hamiltonian at large $U$ which causes the function to increase linearly with $U/8V$.

The next two coefficients $K$ and $M$ are quite similar in the sense that both of them are zero in the first order of the Schrieffer and Wolff transformation. However, $M$ remains zero even after summing to infinite order as shown previously, while $K$ acquires an infinite order singular point at $U = 0$. In this case when $R = S = 0$, the singularity in $K$ is removed as shown in figure 7.14 with the solid line. It starts off from a constant $-1$ at $U = 0$ and increases to zero over $U$, which is identical to the coefficient $G^{(2)} / V$ shown earlier. This is true however only in the symmetric case and in general $K$ behaves very differently to $G$.

$$K^{(2)} = G^{(2)} \quad (7.71)$$

The coefficient $M$ is normally zero as in the result of Schrieffer and Wolff, and
7.3. Elimination of higher order hybridization terms

Figure 7.13: Renormalized coefficient $I/V$ vs $U/8V$, with $t = 0, \mu_c = 0$ and $V = 1$ in the symmetric case when $R = S = 0$ (solid line) and $R = S = Q = T = 0$ (long dashed). The insert shows the same function vs. $8V/U$ when $U = 6$. Grey short dashed is the Schrieffer and Wolff’s result and grey long dashed is our 1D result.

also later in this chapter when all $R, S, T$ and $Q$ are set to zero. The non-zero value of $M$ in this case verifies the imbalance between conduction and localized electrons introduced by the biased condition set in the transformation, to remove hybridization of forms associated only with coefficients $R$ and $S$. In fact the coefficient decreases as $U$ increases. This indicates that the population of the localized electrons outweighs that of the conduction electrons, and is accompanied by the charge interaction with many other electrons, as described by the term in the Hamiltonian Eq. (6.10) associated with the coefficient $M$. This indicates the density of electron is much higher and the interaction between them is more complex.

\[
M^{(2)} = -\frac{(2t - \mu_c - \epsilon_f - U)}{4}(1 - \cos 2\sqrt{2}V\alpha) - \sqrt{2}V \sin 2\sqrt{2}V\alpha
+ \frac{2t - \mu_c - \epsilon_f - U}{2}(1 - \cos 2V\alpha) + V \sin 2V\alpha + G^{(2)} \tag{7.72}
\]

The next two coefficients are from the even order transformation, which still belongs to the hybridization terms. They are however much smaller than the co-
efficients $R$ and $S$, and not as significant. As shown in figure 7.16 and 7.17, the values of these coefficients are close to zero, if not identically zero. At smaller values of $U$, the coefficient reduces but does not vanish. This shows that hybridization is somewhat still in the system but requires the participation of many more electrons.

The formulation of both $T$ and $Q$ are shown below:

\[
T^{(2)} = -\sin 2\sqrt{2}V\alpha + \sin \sqrt{2}V\alpha - \frac{2t - \mu_c - \epsilon_f}{2} \sin 2V\beta \\
- V\cos 2V\beta + V\cos \sqrt{2}V\alpha
\]

\[
Q^{(2)} = -\frac{(2t - \mu_c - \epsilon_f - U)}{2\sqrt{2}} \sin 2\sqrt{2}V\alpha - V\cos 2\sqrt{2}V\alpha \\
- V\cos \sqrt{2}V\alpha + \frac{2t - \mu_c - \epsilon_f - U}{2} \sin 2V\alpha \\
+ \frac{2t - \mu_c - \epsilon_f}{2} \sin 2V\beta + V\cos 2V\alpha + V\cos 2V\beta.
\]
7.3. Elimination of higher order hybridization terms

Figure 7.15: Coefficient $M$ vs $U$, with $t = 0, \mu_c = 0$ and $V = 1$ in the symmetric case when $R = S = 0$ (solid line) and $R = S = Q = T = 0$ (long dashed). The insert shows the same function vs. $8V/U$ when $U = 6$.

One interesting point we would like to discuss is the non-zero value of the $M$ coefficient and the shifting of coefficients from either the upper bound or the lower bound, due to the non-symmetrical condition in the transformation. When the aim of the transformation is to remove certain parts of the hybridization, i.e., only those associated with $R$ and $S$, the Hamiltonian is 'rotated' by the transformation in such a way to serve for that purpose. However, in this case when the terms we want to remove are $f_{i,\sigma}^\dagger c_{i,\sigma} + c_{i,\sigma}^\dagger f_{i,\sigma}$ and $n_{i,-\sigma}^e (f_{i,\sigma}^\dagger c_{i,\sigma} + c_{i,\sigma}^\dagger f_{i,\sigma})$, the transformation creates an imbalance between the localized and the conduction electrons in the hybridization, as terms of the form: $n_{i,-\sigma}^e (f_{i,\sigma}^\dagger c_{i,\sigma} + c_{i,\sigma}^\dagger f_{i,\sigma}) + n_{i,-\sigma}^e (f_{i,\sigma}^\dagger c_{i,\sigma} + c_{i,\sigma}^\dagger f_{i,\sigma})$ are still non-zero. In other words, the transformation suppresses hybridization processes of simple form, but allows those of more complex forms to grow. This is the reason why setting $A = -Z$ does not make $S$ to zero since putting $A = -Z$ forces the terms in $S$ to become $(1 - n_{i,-\sigma}^e) f_{i,\sigma}^\dagger c_{i,\sigma} - c_{i,\sigma}^\dagger f_{i,\sigma}$ and puts a preference on the unoccupied localized electrons in the hybridization. If we put all $R, S, T$ and $Q$ coefficients to
The Three Dimensional Periodic Anderson Model

Figure 7.16: Coefficient $T$ vs $U$, with $t = 0, \mu_c = 0$ and $V = \pm$ in the symmetric case when $R = S = 0$ (solid line) and $R = S = Q = T = 0$ (long dashed). The insert shows the same function vs. $8V/U$ when $U = 6$.

zero as shown in the next section, no preference is made towards either conduction or localized electrons and the transformation treats the localized electrons and the conduction electrons on the same footing.

In fact this imbalance between conduction and localized electrons when $R = S = 0$ is more obvious in the asymmetric case as shown in Appendix H. We will skip the asymmetric case here and defer the discussion until the end of the next section.

The idea of removing the two hybridization terms associated with $R$ and $S$ had also been pursued by Gilles Poirot [74] using a completely different method. His approach started with a similar canonical transformation as ours, but formulated in a more generic Hilbert space. The hybridization terms generated from the transformation, corresponding to our $R$ and $S$ terms, were set to zero using renormalized values of $f_1$ and $f_2$, the counterparts of our $A$ and $Z$. Poirot's method of evaluating the transformation is, however, very different to ours. Instead of using a recursive expansion of commutation $J$, Poirot used the equation of motion from the
7.3. Elimination of higher order hybridization terms

Figure 7.17: Coefficient $Q$ vs $U$, with $t = 0, \mu_c = 0$ and $V = 1$ in the symmetric case when $R = S = 0$ (solid line) and $R = S = Q = T = 0$ (long dashed). The insert shows the same function vs. $8V/U$ when $U = 6$.

Hamiltonian

$$\mathcal{O}(t) = e^{it\Gamma} \mathcal{O} e^{-it\Gamma}$$

(7.75)

to solve for $\mathcal{O}(t)$ from

$$\frac{d}{dt} \mathcal{O}(t) = e^{it\Gamma} [\Gamma, \mathcal{O}] e^{-it\Gamma} = [\Gamma, \mathcal{O}](t) \quad \mathcal{O}(t = 0) = \mathcal{O};$$

(7.76)

and assumed the $n$th commutation of the Hamiltonian with respect to the parameter $\Gamma$ is closed for some large value of $n$.

$$[\Gamma, \mathcal{O}_i] = \sum_j M_{ij} \mathcal{O}_j \quad \text{where} \quad \mathcal{O}_i = [\Gamma, \mathcal{O}_{i-1}, \ldots, \mathcal{O}];$$

(7.77)

with $i \leq n$. The result is then given by the integration of a first order linear equation.

$$\frac{dX_t}{dt} = MX_t \quad \text{with} \quad X_{t=0} = (\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_n)$$

(7.78)

Due to the large value of $n$, the set of differential flow equations needed to be solved is enormous which presents the difficulty of this approach. Poirot managed to
solve it using computer algebra and confirmed the anti-ferromagnetic nature of the Anderson Model in the local moment regime and the intermediate valence regime. Their result is consistent with our calculation using canonical transformation in this section.

7.4 Complete elimination of hybridization terms

The formulation of the last section still suffers from one drawback in which some hybridization terms of more complex form, such as those with coefficients $T$ and $Q$ in Eq. (6.55), are not identically zero. Although their contributions are probably very small and their presence would not have a substantial impact on the characteristics of the system, it is nevertheless not known to us until we have them calculated. If we do have them calculated, we could also in principle set these coefficients to zero in the formulation, similar to what was done to the coefficients $R$ and $S$ of Eq. (6.55) in the last section.

This approach is, however, impossible with the transformation we used in the last section since we had two conditions, namely $R = 0$ and $S = 0$, and precisely two unknowns, $A$ and $Z$. To set four conditions, i.e. $R = 0, S = 0, T = 0$ and $Q = 0$, with only two variables is not possible in general. Nevertheless, we can change the $S$ parameter in the canonical transformation of Eq. (6.1) as follows to accommodate for the inclusion of the hybridization of a more complex form.

$$S = \sum_{i', j', \sigma'} (A + Z n_{j', -\sigma'}^f + X n_{j', -\sigma'}^c + Y n_{j', -\sigma'}^f) (V_{i', j', \sigma'} c_{i', \sigma'}^\dagger f_{j', \sigma'} - V_{i', j', \sigma'}^\dagger c_{i', \sigma'}^\dagger f_{j', \sigma'})$$

in which $A, Z, X$ and $Y$ are unknowns. Using this form of $S$, one can solve for the four unknowns $A, Z, X$ and $Y$ by setting the four coefficients $R, S, T, Q$ to zero. The rest of the algebra is very similar to that presented in the last section, except for the form of $S$ and the corresponding coefficients. In fact, a more general form of the Hamiltonian can be assumed with the use of this new $S$, which provides a more realistic description of the real system. This Hamiltonian, called the extended Anderson lattice model, has all the features of the Anderson Model as well as an extra term for on-site free electron interaction $W c_i^\dagger c_i c_i^\dagger c_i$ which penalizes configurations of two conduction electrons with opposite spin located at the same site of a lattice. Appendix G contains more details of this extended Anderson lattice model, and the results of the canonical transformation applied to this new Hamiltonian. We will however follow the line of the periodic Anderson model and assume the coefficient $W$ to be zero.
Using the result from Appendix G, we can now solve for the parameters $A, Z, X, Y$ from setting $R^{(3)}, S^{(3)}, T^{(3)}$ and $Q^{(3)}$ to zero. A superscript of $(3)$ is added to differentiate these parameters from those of the last section and the one dimensional model. Three more variables $\mu, \nu$ and $\gamma$ are also introduced to simplify the calculation and replace the two counterparts $\alpha$ and $\beta$ in the last section as well as the one dimensional case. They are defined as $\gamma = A, \mu = A + Z$ and $\nu = A + X$. As shown previously, setting $R = 0$ gives the value of $A$, as in Eq. (7.54). This, however, still leaves three variables to be determined from the rest of the conditions. $S$ and $T$ have three variables, $\mu, \nu$ and $\gamma$ which can be solved simultaneously given the value of $\gamma$ from $R = 0$. The two simultaneous equations $S = 0$ and $T = 0$ after some algebraic manipulation are shown below, to solve for $\mu$ and $\nu$.

\[
V \cos(V \sqrt{2(\mu^2 + \nu^2)} + \sin(V \sqrt{2(\mu^2 + \nu^2)} \frac{\mu\nu(2(2-\mu_e-\epsilon_f) + W - U)}{\sqrt{2(\mu^2 + \nu^2)}}) = \kappa, \tag{7.80}
\]

\[
V \cos(2V \sqrt{2(\mu^2 + \nu^2)} + \sin(2V \sqrt{2(\mu^2 + \nu^2)} \frac{2(2-\mu_e-\epsilon_f)(\mu^2 - \nu^2) - U\mu^2 - W\nu^2}{2\sqrt{2(\mu^2 + \nu^2)}}) = \kappa. \tag{7.81}
\]

in which $\kappa = \frac{2t-\mu_e-\epsilon_f}{2} \sin(2VA) + V \cos(2VA)$.

We have kept $\kappa$ in the formulation to be generic, but $\kappa$ in fact is zero, as in the case of the last section. This can be seen from the solution $R = 0$ in Eq. (7.54) that

\[
\sin 2V \gamma = \frac{-2V}{\sqrt{4V^2 + \xi^2}}, \quad \cos 2V \gamma = \frac{\xi}{\sqrt{4V^2 + \xi^2}} \tag{7.82}
\]

where $\xi = 2t - \mu_e - \epsilon_f$. Substituting these into the the equation for $\kappa$ proves the zero identity.

To solve for parameters $X$ and $Z$ (or $\mu$ and $\nu$), the two simultaneous Eqs.(7.80) and (7.81) have to be solved first, which is quite a formidable task. We will however defer this later in this section and first look at them in the limit of the symmetric case when $\epsilon_f = -U/2$. This simplifies the two simultaneous equations drastically and leads us directly to the solution of

\[
\mu = -\nu. \tag{7.83}
\]

This is the counterpart of $A = -Z$ in the Schrieffer and Wolff transformation. Using this result in Eq. (7.81) yields

\[
V \cos 4\mu V - \frac{U}{8} \sin 4\mu V = 0 \tag{7.84}
\]

or

\[
\cos 4\mu V = \frac{U/8}{\sqrt{V^2 + (U/8)^2}} \quad \sin 4\mu V = \frac{V}{\sqrt{V^2 + (U/8)^2}} \tag{7.85}
\]
The \( \mu \) and \( \nu \) are therefore continuous variables with no singularities inherited in their behaviour. All coefficients of the transformed Hamiltonian can now be evaluated.

\[
\begin{align*}
J_{sym}^{(3)} &= -\frac{U}{8} + \sqrt{V^2 + (U/8)^2}, \\
P_{sym}^{(3)} &= \frac{U}{16} - \frac{1}{2} \sqrt{V^2 + (U/8)^2}, \\
G_{sym}^{(3)} &= \frac{U}{4} (1 - \cos 2V\gamma + V \sin 2V\gamma), \\
K_{sym}^{(3)} &= \frac{U}{16} + \frac{1}{2} \sqrt{V^2 + (U/8)^2} + G_{sym}^{(3)}, \\
J_{sym}^{(3)} &= -\frac{U}{16} + \frac{1}{2} \sqrt{V^2 + (U/8)^2} - G_{sym}^{(3)}, \\
M_{sym}^{(3)} &= 0; \\
\end{align*}
\]

in which \( \rho = A + X + Y + Z \) is solved from setting \( Q = 0 \) and given by

\[
\rho = \frac{1}{2V} \tan^{-1}\left(\frac{2V}{2t - \mu_c - \epsilon_f - U + W}\right) 
\]

All of the coefficients are plotted in dark long dashed line in the previous section together with the previous results for comparison. Obviously most of the coefficients are now either sitting at the upper bound or the lower bound of the oscillatory results from the infinite sum of the Schrieffer and Wolff transformation. The imbalance treatment between conduction and localized electrons as mentioned in the last section is removed here in this transformation and the coefficients appear to be as expected with particle-hole symmetry.

In the asymmetric case when the number of free variables has increased to three, namely \( \epsilon_f, U \) and \( V \), the solution of the two simultaneous Eqs. (7.80) and (7.80) are very difficult to find analytically. One easier way is to rewrite them into a form that is possible to solve numerically. In doing so, we introduce two variables \( \phi \) and \( \psi \), different to the ones defined in chapter 6,

\[
\begin{align*}
\tan \phi &= \frac{V(\mu + \nu)\sqrt{2(\mu^2 + \nu^2)}}{\mu \nu (2t - \mu_c - \epsilon_f) + W - U),} \\
\tan 2\psi &= \frac{2V(\mu - \nu)\sqrt{2(\mu^2 + \nu^2)}}{(2t - \mu_c - \epsilon_f)(\mu^2 - \nu^2) - U\mu^2 - W\nu^2} \\
\end{align*}
\]

with which the two simultaneous Eqs. (7.80) and (7.80) become

\[
\begin{align*}
\sin(V\sqrt{2(\mu^2 + \nu^2) + \phi}) &= 0, \\
\sin(2V\sqrt{2(\mu^2 + \nu^2) + 2\psi}) &= 0 \\
\end{align*}
\]
7.4. Complete elimination of hybridization terms

One possible common solution of these two equations is \( \phi = -V \sqrt{2(\mu^2 + \nu^2)} \) and \( \phi = \psi + m\pi \) where \( m \) is an arbitrary number. To remove \( m \) in the equations, we use \( \tan \phi = \tan \psi \) and \( \tan 2\psi = \tan 2\phi = \frac{2\tan \phi}{1 - \tan^2 \phi} \). After substituting Eqs. (7.93) and (7.94) to the two simultaneous Eqs. (7.80) and (7.80), we arrive at a fifth order equation,

\[
2V^2 + (2V^2 + (\xi - U)\Gamma)\chi + ((\xi - U)\Gamma - \Gamma^2)\chi^2 \\
+ (\Gamma^2 - (\xi + W)\Gamma)\chi^3 - (2V^2 + (\xi + W)\Gamma)\chi^4 - 2V^2\chi^5 = 0 .
\]  

(7.96)

in which we defined \( \xi = 2t - \mu_c - \epsilon, \Gamma = 2\xi + W - U \) and \( \chi = \nu/\mu \). One important feature out of this fifth order equation is that it has at least one solution for any value of \( U, V \) and \( \epsilon_f \). This means it is always possible to find a value of \( \mu \) and \( \nu \) in such a way that all the hybridization terms of the Hamiltonian are identically zero.

The coefficients \( J, P, G, K, I \) and \( M \) in the specific case of \( U = 8 \) or \( V = 1 \) or \( \epsilon_f = -3 \) based on this solution are plotted in Appendix H using dark long dashed lines. In general, the singularity is removed from all coefficients which become either smoothly continuous or a step like function with a finite change. These step like changes are caused by the restriction of principal domains between \( \pm \pi/2 \) in the \( \arctan \) function, which can however be removed by relaxing the constraint and allowing the \( \arctan \) function to vary freely. More discussion about the asymmetric case can be found in Appendix H.
Conclusion

Many believed that only non-perturbative approaches such as renormalization group theory, scaling theory and numerical simulations would show the correct behaviours of correlated systems. They believed that perturbation theory did not provide the valid result required for correlated systems, due to the divergence or approximation inherent in the formulation. In actual fact, many condensed matter systems are described by different kinds of perturbative expansions. One example of them is the Bogoliubov transformation [19], a canonical transformation which was expanded to the first order of four particle interactions as described in chapter 5. The result of the transformation not only showed an energy gap between the ‘Cooper pair’ state and the ground state, but also provided the density of states for the superconductive pair from the diagonalized Hamiltonian. The significance of this transformation is the physical insight behind the choice of the pair states and the attractive interaction between electrons. It provided evidence of the effectiveness of perturbation, in the subject of canonical transformation, in condensed matters.

Another example is the evaluation of ground state behaviour for the single impurity Anderson model using perturbation theory. Yei Yosida and Akio Yoshimori [96] obtained the same result using perturbation theory as the non-perturbative results from scaling theory. The key to the success of their perturbation is the proper choice of the singlet ground state which yields a phase shift of $\pm \pi/2$ when a conduction electron at the Fermi surface passes by an impurity. The $\pm$ sign in the phase shift is determined by the relative spin of the conduction and localized electrons. Many physical properties, which were found to be divergent logarithmically in other perturbative approaches, behaved normally in their perturbation without any logarithmic dependence. Again this indicated that the proper choice of ground state is paramount to the validity of the result.

A similar success is also found in Haldane’s perturbative approach [41] to the partition functions for the single impurity Anderson model. As we mentioned in chapter 4, Scalapino [77] and Keiter and Kimball [53] investigated the same model
using perturbation theory about the atomic limit and found leading logarithmic terms of the same form as those occurring in the perturbation expansions for the Kondo model. These leading logarithmic terms as shown by Haldane can in fact be cancelled by terms from the expansion that Scalapino, Keiter and Kimball neglected. Once all the relevant terms were considered, Haldane [41] obtained results using perturbation theory that agreed\(^1\) with the numerical results from the renormalization group theory.

All these examples show that a perturbative approach can be a very powerful tool if applied properly, not just for the weakly interacting system, but also for the strongly correlated system. The choice of the ground state and the selection of the relevant terms are as important as the perturbation calculation itself. We have demonstrated in this work that the ground state of the periodic Anderson model is very likely to be antiferromagnetic, from the result of weak hopping limit and the spin interaction of the atomic limit. This also agrees with the well known antiferromagnetic superexchange interactions between the carrier and the localized electrons in the periodic Anderson model. The results from the weak-hopping limit also suggest that the antiferromagnetic nature of the model is true for different impurity densities, or different doping. All of these are however based upon the assumption that the kinetic energy of the conduction electrons is very small.

When the kinetic energy of the conduction electrons is not small, perturbative expansion in the limit of small hybridization should be used instead. From the canonical perturbation theory, the spin coupling between conduction electrons and impurities is found to be antiferromagnetic when the hybridization is small. This coupling however changes to ferromagnetic as the hybridization becomes more significant. Since the perturbation is calculated based on small hybridization, it is not conclusive whether the spin coupling would indeed change sign in this circumstance. The result of the canonical transformation however confirmed that it does.

The canonical transformation we used was the lattice version of the Schrieffer and Wolff transformation [78] for the periodic Anderson model. This approach has been attempted by many [16,55,69,97,99,100], but only to the low order terms as the higher order terms are deemed to be very complex. The low order results, however, suffer from one common drawback which is plagued by the Schrieffer and Wolff transformation [78], namely the restriction of the parameter space imposed by the negligence of higher order hybridizations. In a small region of the parameter space where the hybridization coefficient \(V\) is small, this has no substantial impact on the

\(^1\)Except that the Kondo temperature is a factor of approximately two higher than the numerical result. Haldane believed that this was due to an error in the evaluation of the non-logarithmic constant term of the second-order perturbation [41].
result of the model as the high order terms are relatively insignificant. However in other interesting regimes such as the mixed valence, the next order term is large enough to change the spin coupling between conduction electrons and impurities, causing the system to turn from antiferromagnetic to ferromagnetic for large $V$. When higher order terms are included in the expansion, this spin coupling changes sign again, leading to a oscillatory behaviour with respect to the values of $U$ and $V$. This oscillation was indeed first observed in one of the plots by Hartmann and Reischl [69] when they considered terms from lower order transformations. It was, however, ignored and considered as an artefact of the calculation as it was outside the radius of convergence. The same oscillation can also be found from the results of E. Kolley, W. Kolley and R. Tietz [55].

Although the first order hybridization process was removed from the Hamiltonian by our canonical transformation, high order hybridization processes were introduced. In an attempt to remove one form of these hybridization processes to all orders out of the Hamiltonian by using another transformation, we removed the oscillation from the spin-carrier coefficient. This indicates that the oscillation is a feature of the transformation which 'orientates' the model in the Hilbert space in such a way that the hybridization terms change sign from one order to another. In this respect, we managed to eliminate all forms of hybridizations from the Hamiltonian by choosing an appropriate transformation to 'align' the model in the Hilbert space orthogonal to the hybridization subspace. This however does not mean that the model should be transformed this way in all circumstances. The decision of which transformation should be used, depends on the nature of model. In the example of the high temperature superconductors, charge fluctuation and hole hopping between Copper ions and Oxygen ions are highly probable [51]. One would therefore have to choose the former transformation which could result in some form of hybridization, between the holes in Copper ions and holes in Oxygen ions. Detailed investigation of this has been reported in [24, 38] and we briefly show the outline below.

The Hamiltonian of the two-band Hubbard model in the standard localized limit [51] commonly used to describe the behaviours of holes in the CuO$_2$ planes, is very similar to the periodic Anderson model. The nomenclature of the coefficients is however different. The hybridization coefficient is labelled '$t_{pd}' instead of 'V', to indicate the hopping matrix element between the p-orbital holes of Oxygen ions and the d-orbital holes of Copper ions. The carriers are the p-orbital holes of Oxygen ions and so $p_{i,\sigma}$ is used to replace $c_{i,\sigma}$. The impurity particles are now the d-orbital holes of the Copper ions, labelled $d_{i,\sigma}$ instead of $f_{i,\sigma}$. The Coulomb interaction
between Oxygen holes is denoted $U_d$, instead of just $U$, to differentiate it from the Coulomb interaction $U_p$ between Copper holes. Our results in the context of holes in this Hamiltonian show the following interactions after a canonical transformation

I) an on-site Hubbard term $U_p(n)$;

II) a Kondo type p-d exchange $J(n)$;

III) an intersite Hubbard $U_{pd}(n)$;

IV) a Josephson two-particle p-d intersite tunnelling $T_{pd}(n)$;

V) a higher-order triplet-creating state $M(n)$.

The new on-site Hubbard term after renormalization with the unperturbed Hamiltonian $U_p$, becomes attractive. This directly implies hole attraction between the p-orbital of Oxygen ions which agrees with earlier quantum Monte Carlo studies [37] and perturbation results [97]. The new Coulomb interaction term between hole carriers, arising from the third and higher odd orders indicates the possibility of charge fluctuations in the high order process. This charge fluctuation is completely suppressed in the local moment limit when $U_d \to \infty$ in the first order result. In this limit, the spin fluctuation dominates the interaction as shown in the result of the Schrieffer and Wolff transformation [78]. As the order of the transformation in the result increases, the spin fluctuation begins to subside when $U_d$ decreases while the charge fluctuation $U_p$ starts to build up [24]. It is important to realize that this result is valid for any value of d- and p- holes hopping as the transformation is summed up to infinite orders.

The spin coupling coefficient $J$ from our result is also interesting since the local limit of the two-band Hubbard model suggests a Kondo-type interaction, rather than a $t - J$ model type interaction. This is however supported by other calculations [55,97] and believed to be the result of the unitary transformation. The spin coupling can be ferromagnetic or antiferromagnetic, depending on the parameters $U_d$ and $t_{pd}$, as explained in the results of the transformation.

In summary, we have explored the periodic Anderson model using different perturbative approaches at different limits of the parameter space. Two canonical methods allowed us to transform the model exactly and obtained results that are valid over the whole parameter space of $U$ and $V$. The new results not only shed light on some interesting behaviours of the model, but also found application to other similar models.
APPENDIX A

The Anderson Model

We introduce the Anderson model in this appendix, which was named after P.A. Anderson [6] in 1970 to explain the magnetic local moments in alloys. This model in Hartree-Fock approximation exhibits very interesting behaviour as shown in Anderson's original paper [6]. This paper is however not entirely correct due to the break down of Hartree-Fock assumption in reality. This simple model remained unsolved for many years, and many different techniques had been developed to look for solutions beyond the mean-field theory, such as perturbation theory, poor's man scaling, slave boson theory, 1/N expansion and non-crossing approximation. All of these theories accurately showed certain aspects of the model, but none of them solved the model exactly until 1981 by Wiegmann [88], following the exact solution of the Kondo model published by Andrei [9,10] in 1980. This solution of the Anderson Model is very specific to the single impurity case which cannot be generalized to the lattice model that we are interested here. Nonetheless, the Single Impurity Anderson Model (SIAM) does provide the essential physics of the model and the foundation on which the lattice case can be built.

The original paper from Anderson for the SIAM remains as one of the classics that are still being quoted in publications nowadays. In this appendix, we will see the physics of the SIAM and its magnetic properties in the mean-field theory. Even though the results from the mean-field approximation are incorrect, it still provides a very elegant way of showing how the Hartree-Fock approximation can be used to find the magnetic phase diagram of the model, using Green's function.

A.1 The reason for a new model

P.A. Anderson [6] in 1970 proposed a model in an attempt to explain the magnetic local moment of alloys. This localized moment cannot be explained by the usual one electron theory which was successful in explaining the properties of non-
magnetic materials. The reason for the failure of this theory is that the bandwidth of the polyvalent metals is so large that the energies of the one electron states coincide with the free electron band. One-electron theory therefore does not permit localization of such a state. Even with virtual states, electrons in such a state would eventually decay into the continuum of free electron states. Further, it is difficult to comprehend how in the Hartree-Fock theory the states of opposite spin in ions can be empty while the parallel spin states are full.

To explain this localized moments in metals, Anderson used the same concept as applied to the insulating magnetic materials that the magnetic state is characterized by the exchange Coulomb interaction. He extended the Hartree-Fock fields for electrons of different spins to include not only the exchange integrals, but also the "true Coulomb integrals". Assuming that the localized moment exists, this means that a d-shell state \( \phi_d \) on the impurity atom of spin up is full; of spin down, empty. If we are to include the repulsive energy within the d-shell in our Hartree field, an electron of spin down will see the repulsion of the extra spin-up electron, while the electron of spin up will not, since they can have no exchange self energy. Thus if the unperturbed energy of the spin-up state lies a distance \( E \) below the Fermi surface, the energy of the spin-down localized state will be \(-E + U\), where \( U \) is the repulsive d-d interaction \([6]\).

Thus the effect of these d state electrons is to reduce the total moment of the system, and hence raises the energy level of the overall system. For instance, if the population of d-state electrons changes by \( \delta n \), the energy between the spin-up and the spin-down states will move from \(-E\) up to \(-E + \delta n U\) for the free electron, while for the d state electrons the energy interval will move from \(-E + U\) down to \(-E + U(1 - \delta n)\). When \( \delta n \) increases, the energy difference between the free electron states and the localized states decreases until to the point that the model starts to break down and becomes completely unstable when it is not possible to maintain a localized moment any longer.

Despite its simplicity, the Anderson model can generate quite a complex behaviour which would have been very difficult to explain by other similar models. The next section will follow very closely to the original paper by P.A. Anderson \([6]\) to introduce the basic characteristics of the model. We will first explain how to turn this true repulsive Coulomb interaction concept into a mathematical expression and incorporate it into the Hamiltonian of the system. It is then followed by the basic thermodynamic properties of the model, using mean-field theory.
A. The Anderson Model

A.2 The Hamiltonian of the new model

The Hamiltonian of a system captures the overall energy of its constituents and their interactions. The Hamiltonian of the Anderson Model should therefore include the “true Coulomb” interaction between the localized electrons. It was P.A. Anderson [6] who first proposed that the Hamiltonian of such a system should be written as

\[ H = \sum_{k\sigma} \varepsilon_k c_{k\sigma}^\dagger c_{k\sigma} + \varepsilon_f (f_{\uparrow}^\dagger f_{\uparrow} + f_{\downarrow}^\dagger f_{\downarrow}) + U n_{\uparrow}^f n_{\downarrow}^f + V_{jk}(c_{k\sigma}^\dagger f_{\sigma} + f_{\sigma}^\dagger c_{k\sigma}), \quad (A.1) \]

where \( c_{k\sigma}^\dagger (c_{k\sigma}) \) create (annihilate) a conduction electrons with spin \( \sigma \) and momentum \( k \), while \( f_{\sigma}^\dagger (f_{\sigma}) \) create (annihilate) a d-orbital impurity electrons of spin \( \sigma \). \( n_{\sigma}^f = f_{\sigma}^\dagger f_{\sigma} \) is the number operator of the d-electron, whereas \( \varepsilon_k \) and \( \varepsilon_f \) are the energies of the conduction and the f-orbital electrons, respectively. \( U \) and \( V \) are the on-site Coulomb repulsion of the d-electrons, and the strength of the hybridization between impurity and conduction electrons. Each term of this Hamiltonian will be explained further in the following.

The \( \varepsilon_k \) in the first term is the energy of free conduction electrons with momentum \( k \). This is the kinetic energy of free electrons which also serves as the unperturbed energy of free electrons in second quantization, in which \( \varepsilon_k \) is the energy of the free electron state of momentum \( k \). \( c_{k\sigma}^\dagger c_{k\sigma} \) or \( n_{k\sigma} \) is the free-electron number operator of momentum \( k \) and spin \( \sigma \). The free electrons are normally assumed to come from the s and p orbital electron shells from which quasi-free electrons are good approximations.

The second term represents the energies of the impurity states. The orbital angular momentum of atoms is ignored and the impurity is considered to have only a single energy level which can be occupied by two electrons of opposite spin. This constitutes what is called a non-degenerated Anderson Model. The degenerated Anderson Model contains more than two electrons occupying the same energy level which can however be studied by extending the non-degenerated model as was shown by Anderson [6].

This second term also serves as the unperturbed energy of the f orbital state in the model. In the Anderson paper, it is used to represent the d orbital state instead of the f state, but in a lot of transition metals and semi-conductors, it is believed that the f-electrons have a more prominent contribution to the characteristics of the system than the d state electron. These f states are assumed to have a single non-degenerated energy level which greatly simplifies the calculation. Furthermore, the same principle can be extended to cover the degenerated model. These f state
electrons are assumed orthogonal to the Wannier functions of free electrons which distinguishes the highly localized characteristic of these electrons from the free electrons. In physical fact, a localized d orbital electron has no resemblance to the band state near it in energy.

The third term represents the repulsive energy among the f orbital state electrons. Although there are repulsion between free electrons, the free electrons are still very much extended throughout the structure than the localized f orbital electrons. Besides, free electrons experience much more screening from the nuclei, than the inner shell electrons do. It is hence not surprising that the $U$ value of the f electrons would be much higher than that of the free electrons, and as a result this repulsion dominates the interaction. As mentioned in [6], $U$ is formally the exchange self-energy of the f electrons, as well as the Coulomb interaction. In order words, we can write the exchange and correlation terms as:

$$\frac{1}{2}K(n_\uparrow n_\downarrow + n_\downarrow n_\uparrow) - \frac{1}{2}J'(n_\downarrow n_\downarrow + n_\uparrow n_\uparrow)$$ (A.2)

in which both $K$ and $J$ are equal to $U$.

The fourth term of the Hamiltonian is the virtual exchange interaction between the conduction electrons and the f electrons at the same site. It describes the hopping of electrons from free conduction band to localized f states and vice versa.

The effect of all these interactions on the electrons in the model are quite complex and tracking the motion of each electron in a lattice is almost impossible. Fortunately, only the thermodynamic average is of concern to us, as average behaviour of the system is what one observes, rather than the behaviour of its individual constituents. In this respect, the mean-field theory or the Hartree-Fock Approximation are normally used to determine the overall behaviour of the model, despite the fact that in the Anderson model the result is actually not correct. Nevertheless, the method does provide a valuable insight into the model and a good understanding on the effect of interaction which would have been very involved mathematically. The next section shows in detail how the Hartree-Fock Approximation can be used to get the correlation of electrons, namely the Greens function, and why it fails in the Anderson Model.
A.3 The properties of the model under Hartree Fock approximation

Under the Hartree-Fock approximation, the wave function of the model can be assumed as

$$\phi_0 = \prod_{\epsilon_n < E_F} c_n^\dagger |\phi_{VAC}\rangle$$  \hspace{1cm} (A.3)

where $|\phi_{VAC}\rangle$ is the wave function of the ground state with no electrons in the lattice, and $c_n^\dagger$ is an electron creation operator of energy $\epsilon_n$, given by

$$[H, c_{n\sigma}^\dagger]_{av} \phi_0 = \epsilon_{n\sigma} c_{n\sigma}^\dagger \phi_0$$  \hspace{1cm} (A.4)

In the formulation, the commutator $[\cdots]_{av}$ takes the "average" of the result of the commutation, in the sense that all number operators $c_{n\sigma}^\dagger c_{n\sigma}$ are to be replaced by their average values for the state $\phi_0$. The Coulomb interacting term is therefore replaced by

$$Un_{d,\uparrow}n_{d,\downarrow} \rightarrow Un_{d,\uparrow}\langle d, \downarrow \rangle + U\langle d, \uparrow \rangle n_{d,\downarrow} - U\langle d, \uparrow \rangle \langle d, \downarrow \rangle$$  \hspace{1cm} (A.5)

This approximation has a substantial consequence to the result. First of all, a number operator replaced by its average value becomes a c-number which is commutative to operators of any kind. When an electron is spin commutative, it would have no effect to any other electrons nearby regardless of their spin. In other words, the electrons have lost their spin property and do not follow the Pauli exclusion principle on which the quantum mechanics built. The approximation therefore removes the local magnetic moment interaction among the electrons and assumes that this interaction have no consequence to the overall behaviour of the model.

The averaging effect of the approximation also deprives the model from any local oscillation and local moments as any local fluctuation will be filtered out during the averaging process. It is indeed this very same local effect which is responsible for the new found phenomena such as Kondo resonance. It is however not to say that one should not take any average at all. After all, we normally only observe the average results of any measurements in laboratories. In fact, taking average can still retain the intrinsic interaction between electrons and the effects from the subtle interplay of electrons, but the key is to find what to average and how to average it. For instance, Kadanoff [50] and Anderson [7] used scaling as a form of averaging to remove high frequency oscillations, by setting a high energy cut-off in the calculation. Wilson [91] exploited the idea of renormalization from high energy physics, and absorbed the unwanted divergence into the basic property of the model.
The averaging in our case is however of the simplest kind and as a result many novel characteristics are lost.

The approximation that Anderson [6] introduced for the SIAM is basically transforming the model into a non-interacting model with an effective d orbital eigenenergy being \( \epsilon_{d,\sigma} = \epsilon_d + U\langle d, -\sigma \rangle \). Without the on-site interaction between localized electrons, the model can be solved easily and the solution found. However, this result no longer reflects the interaction of the electrons in the model.

Let us first resolve \( c_n^\dagger \) into two components, one in terms of \( c_d^\dagger \) the electron creation operator in momentum space and the other one in terms of \( c_k^\dagger \) the impurity electron creation operator.

\[
f c_{n\sigma}^\dagger = \sum_k (n|k\rangle_\sigma c_{k\sigma}^\dagger + (n|d\rangle_\sigma c_{d\sigma}^\dagger)
\]

Substitute Eq. (A.6) into Eq. (A.4) with \( H \) defined in Eq. (A.1) provides,

\[
[H, c_{n\sigma}^\dagger]_{av} = \sum_{k'k\sigma\sigma'} \epsilon_k (n|k\rangle_\sigma [n_{k\sigma}, c_{k'\sigma'}^\dagger]_{av}
+ \sum_{k'} E(n|d\rangle_\sigma [n_{d\uparrow} + n_{d\downarrow}, c_{d\sigma'}^\dagger]_{av}
+ \sum_{k'} U(n|d\rangle_\sigma [n_{d\uparrow} n_{d\downarrow}, c_{d\sigma'}^\dagger]_{av}
+ \sum_{k'k\sigma\sigma'} V_{dk}[c_{k\sigma}^\dagger c_{d\sigma} + c_{d\sigma}^\dagger c_{k\sigma}, (n|k\rangle_\sigma c_{k'\sigma'}^\dagger + (n|d\rangle_\sigma c_{d\sigma'}^\dagger)]_{av} \tag{A.7}
\]

The \([\cdots]_{av}\) in the expression can be evaluated according to its definition and yields

\[
[n_{k\sigma}, c_{k'\sigma'}^\dagger]_{av} = c_{k\sigma}^\dagger \delta_{kk'} \delta_{\sigma\sigma'} \tag{A.8}
\]

\[
[n_{d\uparrow} + n_{d\downarrow}, c_{d\sigma'}^\dagger]_{av} = c_{d\uparrow}^\dagger + c_{d\downarrow}^\dagger \tag{A.9}
\]

\[
[n_{d\uparrow} n_{d\downarrow}, c_{d\sigma'}^\dagger]_{av} = \langle n_{d\uparrow}\rangle c_{d\downarrow}^\dagger + \langle n_{d\downarrow}\rangle c_{d\uparrow}^\dagger \tag{A.10}
\]

\[
[c_{k\sigma}^\dagger c_{d\sigma} + c_{d\sigma}^\dagger c_{k\sigma}, (n|k\rangle_\sigma c_{k'\sigma'}^\dagger + (n|d\rangle_\sigma c_{d\sigma'}^\dagger)] = (n|k\rangle_\sigma c_{d\sigma}^\dagger + (n|d\rangle_\sigma c_{k\sigma}^\dagger) \tag{A.11}
\]

Hence, Eq. (A.7) can be written as

\[
[H, c_{n\sigma}^\dagger]_{av} = \epsilon_k (n|k\rangle_\sigma c_{k\sigma}^\dagger + E(n|d\rangle_\sigma c_{d\sigma}^\dagger + U(n|d\rangle_\sigma c_{d\sigma}^\dagger) + \sum_k V_{dk}(n|k\rangle_\sigma c_{d\sigma}^\dagger + V_{dk}(n|d\rangle_\sigma c_{k\sigma}^\dagger)
= \sum_k (n|k\rangle_\sigma (\epsilon_k c_{k\sigma}^\dagger + V_{dk} c_{d\sigma}^\dagger) + (n|d\rangle_\sigma ((E + U\langle d, -\sigma \rangle) c_{d\sigma}^\dagger + \sum_k V_{dk} c_{k\sigma}^\dagger)
\]

\( (A.12) \)
In fact, we can also evaluate \([H, c_{n\sigma}^\dagger]_{av}\) from substituting Eq. (A.6) into Eq. (A.4)

\[
[H, c_{n\sigma}^\dagger]_{av} \phi_0 = \epsilon_{n\sigma} \left( \sum_k (n|k)_{\sigma} c_{k\sigma}^\dagger (n|d)_{\sigma} c_{d\sigma}^\dagger \right)
\]  

(A.13)

Comparing Eqs. (A.12) and (A.13) by equating the coefficients of \(c_{k\sigma}^\dagger\) and \(c_{d\sigma}^\dagger\), we obtain the coefficients \((n|k)_{\sigma}\) and \((n|d)_{\sigma}\) as

\[
\begin{align*}
\epsilon_{n\sigma}(n|k)_{\sigma} &= (\epsilon_k (n|k)_{\sigma} + V_{dk} (n|d)_{\sigma}) c_{k\sigma}^\dagger \\
\epsilon_{n\sigma}(n|d)_{\sigma} &= (E + U\langle n_{k,-\sigma}\rangle)(n|d)_{\sigma} + \sum_k V_{dk} (n|k)_{\sigma}
\end{align*}
\]  

(A.14)  

(A.15)

These give us the coefficients of the transformation to convert from the \(k\)-momentum space to the \(n\)-number space for the averaged electron operators. The next task is to find the correlation function or the Green’s function of these electron operators from which thermodynamic properties can be derived.

### A.3.1 The Green’s functions solution

The general Green’s function of electrons in a system of Hamiltonian \(H\) is given by

\[
G(\epsilon + is) = \frac{1}{\epsilon + is - H}
\]  

(A.16)

There are however two different kinds of electrons in the Anderson Model which give rise to four different Green’s functions \(G_{kk}, G_{kd}, G_{dk}, G_{dd}\). These are basically projections of the exact Green’s function \(G\) onto the combinations of free electron states and impurity states. The above Green’s function can hence be generalized as

\[
\sum_{\nu}(\epsilon + is - H)_{\mu\nu} G_{\nu k} = \delta_{\mu k}
\]  

(A.17)

in which \(\mu, \kappa, \nu\) can be either \(k\) or \(d\) state. Writing this function out explicitly in terms of \(k\) and \(d\) and using

\[
\begin{align*}
\epsilon &= \epsilon + is \\
E_{\sigma} &= E + U\langle n_{k,-\sigma}\rangle
\end{align*}
\]  

(A.18)  

(A.19)

we get

\[
\begin{align*}
(\epsilon - H)_{dk} G_{kd}^\sigma + (\epsilon - H)_{dd} G_{dd}^\sigma &= \delta_{dd} \\
(\epsilon - H)_{dk} G_{k'd}^\sigma + (\epsilon - H)_{dd} G_{d'd}^\sigma &= \delta_{dk} \\
(\epsilon - H)_{kk'} G_{k'd}^\sigma + (\epsilon - H)_{kd} G_{d'd}^\sigma &= \delta_{kd} \\
(\epsilon - H)_{kk'} G_{kk''}^\sigma + (\epsilon - H)_{kd} G_{kd''}^\sigma &= \delta_{kk''}
\end{align*}
\]  

(A.20)  

(A.21)  

(A.22)  

(A.23)
A.3. The properties of the model under Hartree Fock approximation

Using the Hamiltonian in Eq. (A.1) to evaluate the coefficients of each Green's function,

\[
\begin{align*}
(\varepsilon - H)_{dk} &= \langle d | \varepsilon - H | k \rangle = -\langle d | H | k \rangle = -V_{dk} \tag{A.24} \\
(\varepsilon - H)_{dd} &= \langle d | \varepsilon - H | d \rangle = \varepsilon - \langle d | H | d \rangle = \varepsilon - (E + U \langle n_{d,-}\sigma \rangle) \tag{A.25} \\
(\varepsilon - H)_{kd} &= \langle k | \varepsilon - H | d \rangle = -\langle k | H | d \rangle = -V_{dk} \tag{A.26} \\
(\varepsilon - H)_{kk'} &= \langle k | \varepsilon - H | k' \rangle = \varepsilon - \langle d | H | k \rangle = \varepsilon \delta_{kk'} - \varepsilon_{k'} \delta_{kk'} \tag{A.27}
\end{align*}
\]

Substituting these coefficients back to the four equations of Green's function,

\[
\begin{align*}
(\varepsilon - E_d)G_{dd}^\sigma - \sum_k V_{dk}G_{kd}^\sigma &= 1 \tag{A.28} \\
(\varepsilon - E_d)G_{dk}^\sigma - \sum_{k'} V_{dk}G_{kk'}^\sigma &= 0 \tag{A.29} \\
(\varepsilon - \varepsilon_k)G_{kd}^\sigma - V_{dk}G_{dd}^\sigma &= 0 \tag{A.30} \\
(\varepsilon - \varepsilon_{k'})G_{k'k}^\sigma - V_{dk}G_{dd}^\sigma &= \delta_{kk'} \tag{A.31}
\end{align*}
\]

From the first and the third of these equations, we can eliminate \(G_{kd}^\sigma\) and find out \(G_{dd}^\sigma\).

\[
G_{dd}^\sigma = \frac{1}{\varepsilon - E_d - \sum_k \frac{|V_{dk}|^2}{\varepsilon - \varepsilon_k}} \tag{A.32}
\]

in which the sum in the denominator can further be evaluated as follow

\[
\sum_k \frac{|V_{dk}|^2}{\varepsilon - \varepsilon_k} = \lim_{s \to 0} \sum_k \frac{|V_{dk}|^2}{\varepsilon + is - \varepsilon_k} = \lim_{s \to 0} \sum_k \frac{\varepsilon - \varepsilon_k - is}{(\varepsilon - \varepsilon_k)^2 + s^2} |V_{dk}|^2 \tag{A.33}
\]

from which the real part is

\[
\text{Re}\{\lim_{s \to 0} \sum_k \frac{\varepsilon - \varepsilon_k}{(\varepsilon - \varepsilon_k)^2 + s^2} |V_{dk}|^2\} = \mathcal{P}\{\sum_k \frac{1}{\varepsilon - \varepsilon_k} |V_{dk}|^2\} = \Delta_{E_d} \tag{A.34}
\]

where \(\Delta_{E_d}\) is defined as above and \(\mathcal{P}\) takes the principal value of the argument. The corresponding imaginary part is

\[
\text{Im}\{\lim_{s \to 0} \sum_k \frac{-is}{(\varepsilon - \varepsilon_k)^2 + s^2} |V_{dk}|^2\} = -i\pi \delta(\varepsilon - \varepsilon_k) |V_{dk}|^2 \tag{A.35}
\]

Therefore, the impurity Green's function is

\[
G_{dd}^\sigma = \frac{1}{\varepsilon - E_d - \Delta_{E_d} + i\pi \delta(\varepsilon - \varepsilon_k) |V_{dk}|^2} \tag{A.36}
\]
This result shows that the effective energy is shifted by $\Delta_{E_d}$ due to the presence of the hybridization. The imaginary part in the denominator also indicates the presence of virtual state of energy $i\Delta$ where $\Delta = \pi \delta(\varepsilon - \varepsilon_k)|V_{dk}|^2$. This virtual state is not a bound state and the electrons would not gain any energy in that state. However, electrons do tend to stay in this state for a reasonable period of time before hopping off, resulting in localization to a certain degree. Virtual states can also be found from other systems from another approach as shown by P.A. Wolff [92].

The impurity density distribution can be calculated from the following equation:

$$\rho_{do}(\varepsilon) = \frac{1}{\pi} \sum_n |(d|n)\sigma| ^2 \lim_{s \to 0} \frac{s}{s^2 + (\varepsilon - \varepsilon_{n\sigma})^2}$$

$$= -\frac{1}{\pi} \text{Im}(G^\sigma_{dd}(\varepsilon))$$ (A.37)

in which $G^\sigma_{dd}$ has been evaluated in Eq. (A.36). Note that $\Delta_{E_d}$ is a small real variable which varies the energy level of the system slightly. Since we are only interested in the virtual energy state, it is always possible to absorb the constant $\Delta_{E_d}$ into the variable $\varepsilon$, or even neglect the slight change of energy level.

$$G^\sigma_{dd} = \frac{1}{\varepsilon - E_{\sigma} - \Delta_{E_d} + i\Delta}$$

$$\approx \frac{1}{\varepsilon - E_{\sigma} + i\Delta}$$

$$= \frac{\Delta}{(\varepsilon - E_{\sigma})^2 + \Delta^2} - i \frac{(\varepsilon - E_{\sigma})^2 + \Delta^2}{(\varepsilon - E_{\sigma})^2 + \Delta^2}$$ (A.38)

Therefore,

$$\rho_{do}(\varepsilon) = \frac{\Delta}{\pi (\varepsilon - E_{\sigma})^2 + \Delta^2}$$ (A.39)

Figure A.1 shows the density of this state distribution. We now turn to the Green’s function for the conduction electrons using Eq. (A.23).

$$G^\sigma_{kk'} = \frac{V_{kk'}}{\varepsilon - \varepsilon_{k'}} G^\sigma_{dk} \quad \text{for} (k \neq k')$$ (A.40)

Substituting this back into Eq. (A.21), we have

$$(\varepsilon - E_d)G^\sigma_{dk} - \sum_{k' \neq k} V_{dk'} \frac{V_{k'd}}{\varepsilon - \varepsilon_{k'}} G^\sigma_{dk} = V_{dk} G^\sigma_{kk'}$$ (A.41)

The summation in this equation is identical to the one we have just calculated, which yields $\Delta_{E_d} - i\pi \Delta \approx -i\pi \Delta$. Hence $G^\sigma_{dk}$ can be written in terms of $G^\sigma_{k_k}$ as follows:

$$(\varepsilon - E_{\sigma} + i\Delta + |V_{dk}|^2)G^\sigma_{dk} = V_{dk} G^\sigma_{kk}$$ (A.42)
The properties of the model under Hartree Fock approximation

The diagonal element of the Green’s function in Eq. (A.23) can hence be evaluated as,

\[(\varepsilon - \varepsilon_k) G_{kk}^\sigma = 1 + V_{dk} G_{dk}^\sigma \]
\[= 1 + \frac{\tilde{V}_{dk}^2 G_{kk}^\sigma}{\varepsilon - E_\sigma + i\Delta + \frac{|V_{dk}|^2}{\varepsilon - \varepsilon_k}} \quad \text{(A.43)}\]

Collecting the Green’s function from both sides,

\[G_{kk}^\sigma = (\varepsilon - \varepsilon_k - \frac{|V_{dk}|^2}{\varepsilon - E_\sigma + i\Delta + \frac{|V_{dk}|^2}{\varepsilon - \varepsilon_k}})^{-1} \quad \text{(A.44)}\]

Let \( x = \frac{1}{\varepsilon - \varepsilon_k} \) and \( \xi = \frac{\varepsilon - E_\sigma + i\Delta}{|V_{dk}|^2} \), we have

\[G_{kk}^\sigma = (x^{-1} - \frac{1}{\xi + x})^{-1} = (\frac{\xi + x - x}{x(\xi + x)})^{-1} = \frac{x(\xi + x)}{\xi} \]
\[= x + \frac{x^2}{\xi} = \frac{1}{\varepsilon - \varepsilon_k} + \frac{|V_{dk}|^2}{\varepsilon - E_\sigma + i\Delta}(\varepsilon + \varepsilon_k)^2 \quad \text{(A.45)}\]
A. The Anderson Model

The imaginary part of the Green’s function therefore yields

\[-\frac{1}{\pi} \text{Im} G_{kk}(\epsilon) \approx \frac{1}{\pi} \frac{1}{(\epsilon - \epsilon_k)^2} \left(\frac{\Delta |V_{dk}|^2}{(\epsilon - \epsilon_k)^2 + \Delta^2}\right)\]

\[= \frac{|V_{dk}|^2}{(\epsilon - \epsilon_k)^2} \rho_{d\sigma}(\epsilon)\]  

(A.46)

in which the result of Eq. (A.39) for the density of the d-orbital state has been used. The pole at \(\epsilon_k\) in the equation indicates that the energy state of the free electron remains intact even with the interaction. However, the peak at \(\epsilon_k\) spreads out to the extent of \(\Delta\) due to the density of state factor \(\rho_{d\sigma}\).

A.3.2 Self consistency conditions for localized moments

The main objective of this section is to find the number of d-orbital state electrons, using the density of state from the last section. In fact, the number of d-state electrons of spin \(\sigma\) can be calculated from integrating the density of state in Eq. (A.39) up to the Fermi energy \(\epsilon_F\).

\[\langle n_{d\sigma} \rangle = \frac{1}{\pi} \int_{-\infty}^{\epsilon_F} \frac{\Delta}{(\epsilon - E_{\sigma})^2 + \Delta^2} \, d\epsilon\]  

(A.47)

If we let \(\epsilon - E_{\sigma} = \Delta \cot \theta\), \(d\epsilon = \Delta (\csc^2 \theta) d\theta\),

\[\langle n_{d\sigma} \rangle = \frac{1}{\pi} \int_{\theta(-\infty)}^{\theta(\epsilon_F)} \frac{\Delta^2 (-\csc^2 \theta) d\theta}{\Delta^2 \cot^2 \theta + \Delta^2} = \frac{1}{\pi} \int_{\theta(-\infty)}^{\theta(\epsilon_F)} -d\theta\]

\[= -\frac{1}{\pi} (\cot^{-1} \frac{\epsilon_k - E_{\sigma}}{\Delta} - 0) = \frac{1}{\pi} \cot^{-1} \frac{E_{\sigma} - \epsilon_k}{\Delta}\]  

(A.48)

This equation looks simpler than what it should have been, since \(E_{\sigma}\) is actually proportional to \(\langle n_{d,-\sigma} \rangle\) as shown in Eq. (A.19). Written in an explicit format, the same equation turns into two simultaneous equations in terms of \(\langle n_{d,\sigma} \rangle\) and \(\langle n_{d,-\sigma} \rangle\)

\[\langle n_{d\tau} \rangle = \frac{1}{\pi} \cot^{-1} \left(\frac{E - \epsilon_F + U\langle n_{d,\tau} \rangle}{\Delta}\right)\]  

(A.49)

\[\langle n_{d\iota} \rangle = \frac{1}{\pi} \cot^{-1} \left(\frac{E - \epsilon_F + U\langle n_{d,\iota} \rangle}{\Delta}\right)\]  

(A.50)

To get a self-consistent solution, we let \(y = U/\Delta\) and \(x = (\epsilon_F - E)/U\). We will also drop the angular bracket for \(n\) as it is obvious from the context that \(n\) is not an operator. The two equations can now be generalized into one.

\[\pi n_{d\sigma} = \cot^{-1} (y(n_{d,-\sigma} - x))\]  

(A.51)

These two equations are very difficult to solve, but we can gain insight from investigating three different limits.
I) Magnetic limit: This limit is defined by \( y \gg 1 \) and \( x \) is not small or near 1. Under these conditions, \( \cot^{-1} \) would be close to either zero or \( \pi \) and hence \( n_{d^\uparrow} \) would be either zero or one. Assuming \( n_{d^\uparrow} \approx 1 \) and \( n_{d\downarrow} \approx 0 \), by Taylor expansions of \( \cot^{-1} \)

\[
\begin{align*}
\pi n_{d^\uparrow} & \approx \frac{\pi - \frac{1}{y(x - n_{d\downarrow})}}{y(x - n_{d\downarrow})} \\
\pi n_{d\downarrow} & \approx \frac{1}{y(n_{d^\uparrow} - x)}
\end{align*}
\]

(A.52)

Hence,

\[
\begin{align*}
n_{d^\uparrow} & = 1 - \frac{1}{\pi y(x - n_{d\downarrow})} = 1 - \frac{1}{\pi y(x - \frac{1}{n_{d^\uparrow} - x})} = 1 - \frac{1}{\pi y - \frac{1}{n_{d^\uparrow} - x}}
\end{align*}
\]

(A.53)

and so,

\[
x(1 - n_{d^\uparrow}) = \frac{1}{\pi y - \frac{1}{x(n_{d^\uparrow} - x)}} = \frac{1}{\pi y(1 - \frac{1}{n_{d^\uparrow} - x})}
\]

(A.54)

for the reason that \( n_{d^\uparrow} \) is assumed very close to one. This is the expression for \( n_{d^\uparrow} \) from the approximation. Another result can be obtained from eliminating \( y \) in the two Eqs. (A.52), which gives

\[
n_{d^\uparrow} = 1 - \frac{n_{d^\uparrow} - x}{x - n_{d\downarrow}} \quad n_{d\downarrow} = 1 - \frac{1 - x}{x} n_{d\downarrow}
\]

(A.55)

with the same assumption that \( n_{d^\uparrow} \approx 1 \) and \( n_{d\downarrow} \approx 0 \). Hence,

\[
x(1 - n_{d^\uparrow}) = (1 - x) n_{d\downarrow}
\]

(A.56)

Using both of these results in Eqs. (A.54,A.56),

\[
n_{d^\uparrow} = 1 - \frac{1}{\pi y(1 - \frac{1}{n_{d^\uparrow} - x})}
\]

(A.57)

\[
n_{d\downarrow} = \frac{1}{(1 - x) \pi y(1 - \frac{1}{n_{d\downarrow} - x})}
\]

(A.58)

and the magnetization number is

\[
m = n_{d^\uparrow} - n_{d\downarrow}
\]

\[
= 1 - \frac{1}{\pi y(1 - \frac{1}{n_{d^\uparrow} - x}) - \frac{1}{(1 - x) \pi y(1 - \frac{1}{n_{d\downarrow} - x})}}
\]

(A.59)
II) **Nonmagnetic limit**: In this limit, the simultaneous solution of Eqs. (A.54) and (A.56) becomes

\[
\cot \pi n = y(n - x)
\]

(A.60)

Since \( n \) is close to half, \( \cot \pi n \approx \pi \left( \frac{1}{2} - n \right) \). We can rewrite Eq. (A.60) in this limit to:

\[
\pi \left( \frac{1}{2} - n \right) = y(n - x)
\]

(A.61)

and hence \( n \) becomes

\[
n = \frac{\frac{\pi}{2} + xy}{\pi + y} = \frac{1}{2} \left( 1 + \frac{2y}{\pi} \right)
\]

(A.62)

Therefore, \( n \) is close to \( \frac{1}{2} \) when \( x \) and \( y \) are small, which justifies the assumption set out at the beginning of this calculation.

The effective energy of the d-orbital state electrons relative to the Fermi level is hence,

\[
F_{eff} = U(n - x) = Un - (\epsilon_F - E) - U \frac{1 + 2xy/\pi - 2(x + xy/\pi)}{2(1 + y/\pi)}
\]

\[
= U \frac{1 - 2x}{2(1 + y/\pi)} = \Delta y \frac{1 - 2x}{2(1 + y/\pi)}
\]

(A.63)

in which the definition of \( y \) has been used. This result is valid only when \( x \sim 1/2 \) and \( y \to 0 \). However, if \( x \to 0 \) or \( x \to 1 \) and \( y \) becomes very large,

\[
\cot \pi n \approx \frac{1}{\pi n}, \quad \text{for } n \to 0
\]

(A.64)

Therefore,

\[
\frac{1}{\pi n} \approx y(n - x) \quad \text{or} \quad 0 = \pi n(yn - yx) - 1 = \pi y(n^2 - nx - \frac{1}{\pi y})
\]

(A.65)

When \( y \neq 0 \),

\[
n = \frac{1}{2} (x \pm \sqrt{x^2 + \frac{4}{\pi y}})
\]

(A.66)

Since \( x \) is small and \( y \) is large, \( n \approx \frac{1}{2}(x + x) = x \to 0 \) or \( n = \frac{1}{2}(x - x) = 0 \). Both of them are consistent with the assumption that \( x \) is small. The results of the magnetic and non-magnetic cases are shown in the phase diagram A.2.
A.3.3 The transition between magnetic and non-magnetic regions

Having found the solution at the two extreme limits of the system, we can now work out the transition between the magnetic and the non-magnetic limits. In the transition region, the number of up spin and down spin d-orbital state electrons are still approximately the same, and hence Eq. (A.60) is still true. If we apply this condition to the two simultaneous Eqs. in (A.50) and use Taylor expansion on both \( n_{d\uparrow} \) and \( n_{d\downarrow} \) to the first order, one of the conditions for the transition region will become apparent. First of all, let us define:

\[
f(n) = \frac{1}{\pi} \cot^{-1} \left( \frac{E - \epsilon_F + Un}{\Delta} \right) = \frac{1}{\pi} \cot^{-1} (-xy - yn) \quad (A.67)
\]

Using \( f(n) \), the two simultaneous equations can be re-written as:

\[
\begin{align*}
n_{d\uparrow} &= f(n_{d\downarrow}), \quad n_{d\downarrow} = f(n_{d\uparrow}) \\
(A.68)
\end{align*}
\]

We first assume that there is a solution at \( n_{d\uparrow} = n_{d\downarrow} = n_c \), and expand \( n_{d\uparrow} \) and \( n_{d\downarrow} \) around this solution point in terms of \( \delta n_{d\uparrow} \) and \( \delta n_{d\downarrow} \). The eventual result will confirm the validity of this assumption.

\[
\begin{align*}
n_c &= f(n_c) \\
n_{d\uparrow} + \delta n_{d\uparrow} &= f(n_c) + \delta n_{d\downarrow} f'(n_c) + \cdots \\
n_{d\downarrow} + \delta n_{d\downarrow} &= f(n_c) + \delta n_{d\uparrow} f'(n_c) + \cdots
\end{align*}
\]

In order words,

\[
\begin{align*}
\delta n_{d\uparrow} &\approx \delta n_{d\downarrow} f'(n_c) \\
\delta n_{d\downarrow} &\approx \delta n_{d\uparrow} f'(n_c)
\end{align*} \quad (A.69)
\]

Hence,

\[
\delta n_{d\uparrow} \approx (f'(n_c))^2 \delta n_{d\uparrow} \quad (A.70)
\]

Since \( \delta n_{d\uparrow} \) is not zero and \( f'(n_c) \) is in general real, it concludes that \( f'(n_c) = 1 \). Differentiating \( f(n) \) at \( n = n_c \) yields

\[
\frac{1}{\pi} \left( \frac{y}{1 + (yn_c - xy)^2} \right) = 1 \quad (A.71)
\]

Together with the first condition in Eq. (A.60) at \( n_{d\uparrow} = n_{d\downarrow} \), we have

\[
y = \frac{\pi}{\sin \pi n_c} \quad (A.72)
\]
Unfortunately, Eqs. (A.60) and (A.72) cannot be solved in a simple form to yield \( y_c \), the value of \( y \) at the solution point, as a function of \( x \). Nevertheless, we can still obtain an approximation after simplifying the two equations to:

\[
y_c = \frac{\pi}{\sin^2 2\pi n_c} = \frac{\cot \pi n_c}{n_c - x} \tag{A.73}
\]

or

\[
\sin 2\pi n_c = 2\pi (n_c - x) \tag{A.74}
\]

When \( x \) is close to zero, \( \sin 2\pi n_c \approx 2\pi n_c \) and hence \( n_c \approx 0 \). We can then expand the sine function in Taylor form and obtain

\[
x \approx \frac{(2\pi)^2 n_c^3}{3!} \tag{A.75}
\]

Substituting into Eq. (A.60) and expanding the cosine function in Taylor form

\[
y_c = \frac{\pi}{\sin^2 \pi n_c} = \frac{2\pi}{1 - \cos 2\pi n_c} = \frac{2\pi}{1 - (1 - \frac{(2\pi n_c)^2}{2} + \ldots)}
\approx \frac{4\pi}{(2\pi n_c)^2} \tag{A.76}
\]

Therefore, using Eq. (A.75)

\[
y_c^3 \approx \frac{4\pi}{9x^2} \tag{A.77}
\]

When \( x \approx \frac{1}{2} \), we can expand the sine function around the point \( x = \frac{1}{2} \). Let \( z = x - \frac{1}{2} \). Eq. (A.72) can be re-written as

\[
\sin(2\pi((n_c + \frac{1}{2}) - \frac{1}{2})) = 2\pi(n_c + \frac{1}{2} - (x - \frac{1}{2}))
\]

\[
\sin(2\pi(n_c + \frac{1}{2}) - \pi) = 2\pi(n_c + \frac{1}{2} - z)
\]

\[
- \sin(2\pi(n_c + \frac{1}{2})) = 2\pi((n_c + \frac{1}{2}) - z)
\]

\[
-2\pi(n_c + \frac{1}{2}) - \frac{1}{3!}(2\pi(n_c + \frac{1}{2}))^3 + \cdots = 2\pi((n_c + \frac{1}{2}) - z)
\]

\[
-2 \cdot 2\pi(n_c + \frac{1}{2}) \approx -2\pi z
\]

\[
(n_c + \frac{1}{2}) = \frac{1}{2} z
\]
A.4. Failure of the Approximation

Failure of the Approximation

As mentioned in the beginning, the approximation is in fact not satisfactory because the local moment regime which was defined by \( \langle n_{d_1} \rangle \neq \langle n_{d_2} \rangle \), breaks the in which \( z \) is assumed non-zero. Substituting this into Eq. (A.72),

\[
y_c = \frac{2\pi}{1 - \cos 2\pi n_c} = \frac{2\pi}{1 + \cos(2\pi(n_c + \frac{1}{2}))}
\]

\[
= 2\pi(1 + 1 - \frac{1}{2!}(2\pi(n_c + \frac{1}{2})^2 + \cdots))^{-1}
\]

\[
= \pi(1 - \pi^2(n_c + \frac{1}{2})^2 + \cdots)
\]

\[
= \pi - \pi^3 z^2 / 4
\]

\[
= \pi - \frac{1}{4} \pi^3 (x - \frac{1}{2})^2
\]

The result is summarized in the two plots of figure A.3.

Figure A.2: The phase diagram of the single impurity Anderson model derived under the mean field theory.

Figure A.3: The average number of impurity as a function of \( \pi/y = \pi\Delta/U \). \( \langle n_+ \rangle - \langle n_- \rangle \) provides the overall number of spins which vanishes at the transition \( y_c \) of the triple point. (a) shows when \( x = 1/2 \), (b) when \( x = 1/4 \).
A. The Anderson Model

local symmetry of the Hamiltonian. Even in the paramagnetic regime in which \( \langle n_d, \uparrow \rangle = \langle n_d, \downarrow \rangle \), the criterion used for the solution is still not valid in general. The experimental criterion for a local moment in the result is that a Curie-Weiss term must be found in the impurity susceptibility, and this can depend on the temperature range being considered which is however absent in the Hartree-Fock approximation. The Random Phase Approximation (RPA) had been used on the Anderson Model but the solution became unstable at some finite value of \( U \).

Perturbation became one of the main tools to find an approximated solution for the Anderson Model. The perturbation required can be a direct expansion of the partition function in a power series of \( U \) [95] and \( V \) [41]. The perturbation in a power series of \( U \) in [95] used the non-magnetic ground state of the unperturbed Hamiltonian including the s-d mixing term and no correlation, to expand the partition function. In the symmetric case where \( \epsilon_d = -U/2 \) and \( \epsilon_d = 0 \), Yamada and Yosida [95] showed that each term of the perturbation expansion can be described by the imaginary-time integral of the fourth power of the Pfaffian constructed from the d-electron Green’s functions. Using an asymptotic form of a Green’s function, an one particle Hamiltonian of the Anderson Model was solved. The results agreed with the numerical work [57], and with the phenomenological Fermi liquid theory of Nozieres [72], when \( U \) is small relative to the resonance width \( \Delta = \pi \rho |V^2| \) where \( \rho \) is the density of state of the conduction electrons. However, when the system is not symmetrical and \( \epsilon_d \neq 0 \), the determinant expressions in the result become very complicated. The convergence of the determinant expansions also becomes rather poor when the model is highly asymmetrical with large Coulomb interaction between the localized electrons [47]. In fact, the ground state energy of the analysis becomes highly unstable when the higher order perturbation terms becomes significant as \( U \) gets larger. In this case, expanding the Hamiltonian in the power series of \( V \) will provide a much better convergence. The perturbation with respect to a small value of \( V \) is however very difficult to calculate as the determinant expressions become very complex. Another approach is taken by Haldane in [41] to expand the partition function in power series of \( V \), as already explained in detail in chapter 4.

The scaling method was also used to solve the Anderson model in very similar fashion to that of the Kondo model. The scaling approach basically involves two steps.

1) The first step is to divide the energy spectrum into two portions, one containing most of the lower energy of the system while the other one has the highest energy. The high energy portion of the energy spectrum is normally very small and can be removed without causing a substantial impact on the
system.

II) The second step is to change the energy scale of the system to span the entire energy spectrum by the cut-off version of the new spectrum. Since the divergence associated with the breakdown of perturbation theory depends on the conduction band width $D$ in this case, the scaling is best applied on this band width. Electrons fallen into the states corresponding to the higher energy portion of the band width is removed before re-scaling the band width. The fact that the divergence is logarithmic in $D$, not $1/D$ or $1/D^2$, clearly indicates that the higher energy excitations are important and have to be taken into account, when $D \to \infty$. The essence of the scaling approach is that these higher energy excitations can be absorbed into a renormalization factor of the corresponding interaction. The logarithmic terms show only in the parameters of the effective Hamiltonian for the calculation of the low lying levels.

The scaling idea is powerful, but it does not provide a solution for the thermodynamics properties of the Anderson Models in the local moment regime in which $\epsilon_d + U \gg \epsilon_F$ and $\epsilon_d \ll \epsilon_F$, except in the case of ferromagnetic coupling. This is because the scaling approach is still perturbative and would break down at the singular point for an antiferromagnetic interaction between local moment and conduction electrons. Nevertheless they have been useful in extending the leading order perturbation results. More importantly they have provided conceptual framework, clarifying what is required at the next stage of non-perturbation scheme for calculating the excitations beyond the reach of perturbative calculation. More detail for the SIAM is included in Ref. [43].
APPENDIX B

Full results of 2nd and 3rd order commutations

We show in this appendix the full result of the second order and the third order commutation in three dimensional space, before index contraction. The second order result is:

\[
[S, [S, H_V]] = \sum_{ijkl} -4AAV_{kl}V_{ij}V_{kl}^\dagger c_{i+\sigma}^\dagger f_{j+\sigma} + 4AAV_{ji}V_{kl}V_{ik}V_{ij}^\dagger c_{i+\sigma} f_{j+\sigma}
\]

\[+2(AZ - ZZ)V_{ik}V_{ij}V_{kl}^\dagger c_{j+\sigma}^\dagger c_{i+\sigma} c_{i+\sigma} f_{j+\sigma} + 2(AZ - ZZ)V_{ij}V_{ik}V_{kl}V_{ij}^\dagger c_{i+\sigma} c_{i+\sigma} c_{i+\sigma} f_{j+\sigma}
\]

\[-(5AZ + ZZ)V_{il}V_{lj}V_{kl}^\dagger f_{l-i} f_{i-\sigma} f_{i+\sigma} + 3AZV_{lj}V_{kl}V_{il}V_{ij}^\dagger c_{j+\sigma} c_{j+\sigma} f_{i+\sigma}
\]

\[-3(3ZZV_{il}V_{lj}^\dagger c_{j+\sigma} c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma} + 3AZV_{lj}V_{kl}V_{il}V_{ij}^\dagger c_{j+\sigma} c_{j+\sigma} c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma}
\]

\[+3(3ZZV_{il}V_{lj}^\dagger c_{j+\sigma} c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma} + 3AZV_{lj}V_{kl}V_{il}V_{ij}^\dagger c_{j+\sigma} c_{j+\sigma} c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma}
\]

\[-2AZV_{lj}V_{kl}V_{il}^\dagger c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma} - (5AZ + ZZ)V_{ik}V_{jl}V_{kl}^\dagger c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma}
\]

\[-2AZV_{lj}V_{kl}V_{il}^\dagger c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma} - (5AZ + ZZ)V_{ik}V_{jl}V_{kl}^\dagger c_{j+\sigma} f_{l-i} f_{i-\sigma} f_{i+\sigma}
\]

The general third order commutation result is:

\[
[S, [S, [S, H_V]]] = \sum_{ijklm} +8AAAV_{km}V_{lj}V_{kl}V_{lm}^\dagger c_{i+\sigma} c_{j+\sigma}
\]

\[-8AAAV_{lm}V_{jk}V_{jl}^\dagger V_{km}^\dagger c_{i+\sigma} f_{j+\sigma} + [(3AZ - AAZ)V_{lm}V_{jk}V_{jl}V_{km}^\dagger + 4(AAZ - AZZ)V_{mk}V_{ml}V_{mj}^\dagger V_{km}^\dagger c_{j+\sigma} c_{l+\sigma} c_{k+\sigma} c_{k+\sigma}
\]

\[+(2ZZZ - 8AZZ)V_{ij}V_{ij}^\dagger V_{ij}V_{ij}^\dagger c_{k+\sigma} c_{k+\sigma} f_{i+\sigma} f_{i+\sigma}
\]

232
\[-12AAZV_{im}V_{ij}^\dagger V_{ik}^\dagger V_{lm}^\dagger c_{j+\sigma}^i f_{i+\sigma} f_{i-\sigma} \]
\[-[(12AAZ + 4AZZ)V_{im}V_{ij}^\dagger V_{ik}^\dagger V_{lm}^\dagger + (12AAZ + 4AZZ)V_{ij}V_{im}^\dagger V_{ik}^\dagger + 12(AZZ - ZZZ)V_{ij}V_{il}^\dagger V_{ik}^\dagger c_{j+\sigma}^i f_{i+\sigma} f_{i-\sigma} \]
\[+[(12AAZ + 4AZZ)V_{ij}V_{lm}^\dagger V_{im}^\dagger + (16AZZ + 8ZZZ)V_{ij}V_{il}^\dagger V_{ik}^\dagger + (4AZZ + 12AAZ)V_{im}V_{ij}^\dagger V_{lk}^\dagger c_{j+\sigma}^i f_{i+\sigma} f_{i-\sigma} \]
\[+[(4AZZ + 12AAZ)V_{ik}V_{lm}^\dagger V_{il}^\dagger c_{j+\sigma}^i f_{i+\sigma} f_{i-\sigma} + 4[(AAZ - AAZ)V_{ij}V_{im}^\dagger V_{ik}^\dagger \]
\[+(AAZ - AAZ)V_{ij}V_{il}^\dagger V_{lm}^\dagger c_{j+\sigma}^i f_{i+\sigma} f_{i-\sigma} \]
\[-[2AZZV_{ij}V_{im}V_{ik}^\dagger V_{lm}^\dagger + (6AZZ + 3ZZZ)V_{ij}V_{ik}^\dagger V_{il}^\dagger c_{j+\sigma} c_{k-\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+12AAZV_{ij}V_{ik}V_{lm}^\dagger V_{il}^\dagger c_{j+\sigma} c_{k-\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[-3AZZV_{ik}V_{ij}^\dagger V_{il}^\dagger V_{j k}^\dagger f_{i+\sigma} f_{j+\sigma} f_{j-\sigma} \]
\[-3AZZV_{ij}V_{lm}^\dagger V_{ik}^\dagger V_{j k}^\dagger f_{i+\sigma} f_{i-\sigma} f_{j-\sigma} \]
\[+2(5AZZ + AAZZ)V_{ij}V_{il}^\dagger V_{ik}^\dagger V_{j k}^\dagger f_{i+\sigma} f_{i-\sigma} f_{j-\sigma} \]
\[+3AZZV_{j m}V_{ik}V_{il}^\dagger f_{j+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+6ZZZV_{ij}V_{ik}^\dagger V_{lm}^\dagger c_{j+\sigma} c_{k+\sigma} f_{i+\sigma} \]
\[-(7AAZV_{ik}V_{ml}V_{j k}^\dagger V_{j m}^\dagger + 6AZZV_{ik}V_{il}^\dagger V_{j k}^\dagger V_{j l}^\dagger c_{j+\sigma} c_{k-\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[-[7AAZV_{jk}V_{ml}V_{ij}^\dagger V_{jm}^\dagger + (8AZZ + 2ZZZ)V_{ij}V_{jm}^\dagger V_{jl}^\dagger c_{j+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+2AZZ + 10AAZ)V_{im}V_{kl}^\dagger V_{jm}^\dagger V_{il}^\dagger V_{jk}^\dagger f_{i+\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} \]
\[+(4AZZ + 2ZZZ)V_{il}^\dagger V_{jm}^\dagger V_{ik}^\dagger V_{j k}^\dagger c_{j+\sigma} c_{k+\sigma} c_{m-\sigma} c_{l+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[-(8AZZ + 4ZZZ)V_{ik}V_{il}^\dagger V_{j k}^\dagger V_{jm}^\dagger c_{j+\sigma} c_{k+\sigma} c_{m-\sigma} c_{l+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+(16AZZ + 8ZZZ)V_{ik}V_{il}^\dagger V_{jm}^\dagger c_{j+\sigma} c_{k+\sigma} c_{m-\sigma} c_{l+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+(4AZZ + 8AZZ)V_{ik}V_{il}^\dagger V_{jm}^\dagger f_{i+\sigma} f_{i-\sigma} c_{j+\sigma} c_{k+\sigma} c_{m-\sigma} c_{l+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+(11AZZ + 8ZZZ)V_{ik}V_{il}^\dagger V_{jm}^\dagger f_{i+\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} \]
\[+(4AZZ + 8AZZ)V_{ik}V_{il}^\dagger V_{jm}^\dagger f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} c_{j+\sigma} c_{k+\sigma} c_{m-\sigma} c_{l+\sigma} f_{i+\sigma} f_{i-\sigma} \]
\[+(2AZZ + 8ZZZ)V_{ik}V_{il}^\dagger V_{jm}^\dagger c_{j+\sigma} c_{k+\sigma} c_{m-\sigma} c_{l+\sigma} f_{i+\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} \]
\[-3AZZV_{ij}V_{il}^\dagger V_{jm}^\dagger c_{j+\sigma} c_{k+\sigma} f_{i+\sigma} f_{i-\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} f_{i+\sigma} \]
\[-(3AZZV_{im}V_{ij}^\dagger V_{kl}^\dagger V_{j m}^\dagger + (ZZZ + 5AZZ)V_{jm}V_{ik}^\dagger V_{jl}^\dagger V_{jk}^\dagger c_{j+\sigma} c_{k-\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} \]
\[+(6AZZ + 3ZZZ)V_{jm}V_{ik}^\dagger V_{jl}^\dagger V_{j k}^\dagger c_{j+\sigma} c_{k-\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} \]
\[+(2ZZZ - 5AZZ)V_{jm}V_{ik}^\dagger V_{jl}^\dagger V_{j k}^\dagger f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} \]
\[-(4AZZ + 2ZZZ)V_{ij}V_{ik}^\dagger V_{jl}^\dagger c_{j+\sigma} c_{k-\sigma} c_{m-\sigma} c_{l-\sigma} f_{i+\sigma} f_{i-\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} f_{i+\sigma} \]
\[-3ZZZV_{ij}V_{il}^\dagger c_{j+\sigma} c_{k+\sigma} f_{i+\sigma} f_{i-\sigma} f_{j+\sigma} f_{j-\sigma} f_{i-\sigma} f_{i+\sigma} \]
B. Full results of 2nd and 3rd order commutations

\[-[(ZZZ + 14AZZ)_{ij} V_{ik} V_{im} V_{im}^{\dagger} V_{im}^{\dagger} + (2AZZ - 5ZZZ)_{ij} V_{ik} V_{im} V_{ik} V_{im} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

\[-[(8AZZ + ZZZ)_{ij} V_{ik} V_{im} V_{im}^{\dagger} V_{im}^{\dagger} - 3ZZZ_{ij} V_{ik} V_{im} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

\[-[(8AZZ + 16AZZ)_{ij} V_{im} V_{jl} V_{im}^{\dagger} V_{im}^{\dagger} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

\[-[(16AZZ + 8ZZZ)_{ij} V_{im} V_{jl} V_{im}^{\dagger} V_{im}^{\dagger} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

\[-[(8AZZ + ZZZ)_{ij} V_{km} V_{il} V_{im}^{\dagger} V_{im}^{\dagger} c_{j} c_{j} c_{j} c_{j} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

\[-[(5AZZ + ZZZ)_{ij} V_{im} V_{ik} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

\[-[(3ZZZ - 3AZZ)_{ij} V_{im} V_{ik} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+[(6AZZ)_{ij} V_{im} V_{ik} V_{im}^{\dagger} V_{im}^{\dagger} + (2AZZ - ZZZ)_{ij} V_{ik} V_{im} V_{ik} V_{im} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+[(8AZZ + 16AZZ)_{ij} V_{im} V_{jl} V_{im}^{\dagger} V_{im}^{\dagger} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+[(16AZZ + 5ZZZ)_{ij} V_{im} V_{jl} V_{im}^{\dagger} V_{im}^{\dagger} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+[(2AZZ - 5ZZZ)_{ij} V_{il} V_{im} V_{ik} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

-3ZZZ V_{ij} V_{ik} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+[(AZZ + 2ZZZ)_{ij} V_{im} V_{ik} V_{im}^{\dagger} V_{im}^{\dagger} c_{k+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+6AZZ V_{ij} V_{km} V_{il} V_{im}^{\dagger} V_{im}^{\dagger} c_{j} c_{j} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+9AZZ V_{ij} V_{km} V_{il} V_{im}^{\dagger} V_{im}^{\dagger} c_{j} c_{j} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

-3ZZZ V_{ij} V_{ik} V_{jm}^{\dagger} V_{jm} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

-3AZZ V_{ij} V_{ik} V_{jm}^{\dagger} V_{jm} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+5AZZ V_{ij} V_{ik} V_{jm}^{\dagger} V_{jm} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+5AZZ V_{ij} V_{ik} V_{jm}^{\dagger} V_{jm} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+5AZZ V_{ij} V_{ik} V_{jm}^{\dagger} V_{jm} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}

+5AZZ V_{ij} V_{ik} V_{jm}^{\dagger} V_{jm} c_{l+\sigma} c_{j-\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma} f_{i-\sigma} f_{i+\sigma} f_{j-\sigma} f_{j+\sigma}
Mathematically speaking, we can re-write Eq. (7.2) for example to differentiate the different types of interaction among the neighbouring electrons.

\[
H_1 = \sum_{i,j,k,\sigma} X_{i,j,k,\sigma}
= \sum_{i,\sigma} X_{i,j=i,k=i,j'=i,1,\sigma}
+ \sum_{i,\sigma} (X_{i,j=i,k=i+1,\sigma} + X_{i,j=i+1,k=i} + X_{i,j=i+1,k=i+1})
+ \sum_{i,\sigma} (X_{i,j=i,k=i+2,\sigma} + X_{i,j=i+2,k=i} + X_{i,j=i+2,k=i+2}) + \cdots \tag{C.1}
\]

in which \(X_{i,j,k}\) is a generic expression for the summation in Eq. (7.2) with three different summing indices. Based on this argument, the main contribution comes from the first term of Eq. (C.1) with all three indices being identical. The first correction includes the subsequent terms in the second summation with at least one of the three indices offset by one site. The second correction is similarly considered as a result of the third summation in which the interacting electrons are at least two sites apart, also known as the next nearest neighbours.

In this thesis, we focus on the major contribution of the result from the transformation with all indices being identical, which is believed to account for most of the observable properties of the model.
APPENDIX D

Relevance of order between commutation and diagonalization

In this appendix, we show the sequence of commutation and diagonalization has no effect to the final result. The diagonalization as explained in chapter 7 is essentially the index contraction through delta functions. The commutation operation here is between a Hamiltonian and the parameter $S$ as defined in Eq. (6.1), whereas the index contraction is simply summing the operators with the same indices. Normally to simplify a result from canonical transformation, one would first need to carry out the commutation of the Hamiltonian with $S$, and then apply a delta function with respect to all the summing indices to obtain the first order result. However, if the delta function is imposed on the operands before the commutation is carried out, the computation is dramatically simplified and the result will still remain the same.

To justify this, let us consider the commutation of $R_{i,j,k,l,...}$ with $S_{a,b}$ in which both terms contain operators of conduction and/or impurity electron at sites $i,j,k,l,...$ and $a,b$.

\[
\sum_{a,b,i,...} [R_{i,j,k,l,...}, S_{a,b}] \cdot \delta_{a,b,i,j,k,l,...}
= \sum_{a,b,i,...} (R_{i,j,k,l,...} S_{a,b} - S_{a,b} R_{i,j,k,l,...}) \delta_{a,b,i,j,k,l,...}
= \sum_{a,b,i,...} R_{i,j,k,l,...} \delta_{a,b,i,j,k,l,...} - \sum_{a,b,i,...} S_{a,b} R_{i,j,k,l,...} \delta_{a,b,i,j,k,l,...}
= \sum_{i} R_{i} S_{i} - S_{i} R_{i}
= \sum_{i} [R_{i}, S_{i}] \tag{D.1}
\]
In the next commutation of this result with $S_{a,b}$,

$$\sum_{a,...,i,...} [[R_{i,j,k,l,...}, S_{a,b}], S_{c,d}] \cdot \delta_{a,...,i,...}$$

$$= \sum_{a,...,i,...} ([R_{i,j,k,l,...}, S_{a,b}]S_{c,d} - S_{c,d}[R_{i,j,k,l,...}, S_{a,b}]) \cdot \delta_{a,...,i,...}$$

$$= \sum_{a,...,i,...} [R_{i,j,k,l,...}, S_{a,b}]S_{c,d}\delta_{a,b,c,d,i,j,k,l,...} - \sum_{a,...,i,...} S_{c,d}[R_{i,j,k,l,...}, S_{a,b}]\delta_{a,...,i,...}$$

$$= \sum_{i} [R_{i}, S_{i}] S_{i} - S_{i}[R_{i}, S_{i}]$$

$$= \sum_{i} [[R_{i}, S_{i}], S_{i}] \quad (D.2)$$

Therefore the order of the commutation and the application of the delta function can be changed without affecting the result. Using this approach, we manage to reduce the computational time dramatically and have the high order results produced within a few seconds, using a PC486/100 computer with 16MB of RAM.
APPENDIX E

Computer generated transformation results of low orders

The results of the first eleven order commutation of $H_V$ with $S$ as defined in Eq. (6.9) are collected in the three tables below. The coefficients $J, P, G, M, I, K$ are defined in Eq. (6.10) which is reproduced here for reference.

$$[[S, H_V]]_n = \sum_{i,\sigma} \left[ J_n (c_{i,\sigma} c_{i,-\sigma} f_{i,-\sigma} f_{i,\sigma} - n_{i,\sigma}^\xi n_{i,-\sigma}^f) 
+ P_n (c_{i,\sigma} c_{i,-\sigma} f_{i,\sigma} f_{i,-\sigma} + f_{i,\sigma} f_{i,-\sigma} c_{i,\sigma} c_{i,-\sigma}) + G_n (n_{i,\sigma}^f - n_{i,\sigma}^\xi) + I_n n_{i,-\sigma}^f n_{i,\sigma}^f 
+ M_n n_{i,-\sigma}^f n_{i,-\sigma}^\xi (n_{i,\sigma}^f - n_{i,\sigma}^\xi) + K_n n_{i,-\sigma}^f n_{i,\sigma}^\xi \right],$$

for odd value of $n$. 

239
The table below lists the coefficients $J, P, G$ and $M$ of the generalized odd order Hamiltonian term in Eq. (6.10).

<table>
<thead>
<tr>
<th>Order</th>
<th>Prefactor</th>
<th>Coefficient $J$</th>
<th>Coefficient $P$</th>
<th>Coefficient $G$</th>
<th>Coefficient $M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(1 - \frac{1}{2})V^2$</td>
<td>$+2Z$</td>
<td>$-Z$</td>
<td>$+2A$</td>
<td>null</td>
</tr>
<tr>
<td>2</td>
<td>$(\frac{1}{2} - \frac{1}{3})V^3$</td>
<td>null</td>
<td>null</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>3</td>
<td>$(\frac{1}{3} - \frac{1}{4})V^4$</td>
<td>$-32A^2Z$</td>
<td>$-8A^2Z$</td>
<td>$-16Z^2$</td>
<td>$8Z^2$</td>
</tr>
<tr>
<td>4</td>
<td>$(\frac{1}{4} - \frac{1}{5})V^5$</td>
<td>null</td>
<td>$+224A^4Z$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>5</td>
<td>$(\frac{1}{5} - \frac{1}{6})V^6$</td>
<td>$+100A^4Z^2$</td>
<td>$+448A^3Z^2$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>6</td>
<td>$(\frac{1}{6} - \frac{1}{7})V^7$</td>
<td>null</td>
<td>null</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>7</td>
<td>$(\frac{1}{7} - \frac{1}{8})V^8$</td>
<td>$-396A^6Z$</td>
<td>$-1104A^5Z^2$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>8</td>
<td>$(\frac{1}{8} - \frac{1}{9})V^9$</td>
<td>$+131072A^8$</td>
<td>$+56028A^8Z^2$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>9</td>
<td>$(\frac{1}{9} - \frac{1}{10})V^{10}$</td>
<td>$+524288A^7Z^2$</td>
<td>$+260096A^7Z^3$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>10</td>
<td>$(\frac{1}{10} - \frac{1}{11})V^{11}$</td>
<td>null</td>
<td>$+512A^9$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>11</td>
<td>$(\frac{1}{11} - \frac{1}{12})V^{12}$</td>
<td>null</td>
<td>null</td>
<td>null</td>
<td>null</td>
</tr>
</tbody>
</table>

Table E.1: The coefficients $J, P, G$ and $M$ of the generalized odd order Hamiltonian term in Eq. (6.10).
<table>
<thead>
<tr>
<th>Order</th>
<th>Prefactor</th>
<th>Coefficient $I$</th>
<th>Coefficient $K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(1 - \frac{1}{2}) V^2$</td>
<td>$+2Z$</td>
<td>null</td>
</tr>
<tr>
<td>2</td>
<td>$(\frac{1}{2} - \frac{1}{3}) V^3$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>3</td>
<td>$(\frac{1}{3} - \frac{1}{4}) V^4$</td>
<td>$-28A^2Z$</td>
<td>$-4A^2Z$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-36AZ^2$</td>
<td>$+4AZ^2$</td>
</tr>
<tr>
<td>4</td>
<td>$(\frac{1}{4} - \frac{1}{5}) V^5$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>5</td>
<td>$(\frac{1}{5} - \frac{1}{6}) V^6$</td>
<td>$+336A^4Z$</td>
<td>$+176A^4Z$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+848A^3Z^2$</td>
<td>$+88A^3Z^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+936A^2Z^3$</td>
<td>$-8AZ^4$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+520AZ^4$</td>
<td>$+128Z^7$</td>
</tr>
<tr>
<td>6</td>
<td>$(\frac{1}{6} - \frac{1}{7}) V^7$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>7</td>
<td>$(\frac{1}{7} - \frac{1}{8}) V^8$</td>
<td>$-4544A^6Z$</td>
<td>$-3648A^6Z$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-17088A^5Z^2$</td>
<td>$-7488A^5Z^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-29312A^4Z^3$</td>
<td>$-7552A^4Z^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-29024A^3Z^4$</td>
<td>$-3744A^3Z^4$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-17520A^2Z^5$</td>
<td>$-912A^2Z^5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-6160AZ^6$</td>
<td>$+164AZ^6$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-1024Z^7$</td>
<td>$-1059840A^{10}Z$</td>
</tr>
<tr>
<td>8</td>
<td>$(\frac{1}{8} - \frac{1}{9}) V^9$</td>
<td>$+67840A^8Z$</td>
<td>$63232A^8Z$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+332544A^7Z^2$</td>
<td>$191744A^7Z^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+759936A^6Z^3$</td>
<td>$288640A^6Z^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+1053952A^5Z^4$</td>
<td>$256768A^5Z^4$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+969792A^4Z^5$</td>
<td>$144320A^4Z^5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+607424A^3Z^6$</td>
<td>$47936A^3Z^6$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+254240A^2Z^7$</td>
<td>$7904A^2Z^7$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+65568AZ^8$</td>
<td>$-32AZ^8$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$+8192Z^9$</td>
<td>null</td>
</tr>
<tr>
<td>9</td>
<td>$(\frac{1}{9} - \frac{1}{10}) V^{10}$</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>10</td>
<td>$(\frac{1}{10} - \frac{1}{11}) V^{11}$</td>
<td>$-1059840A^{10}Z$</td>
<td>$-1037312A^{10}Z$</td>
</tr>
<tr>
<td></td>
<td>$(\frac{1}{11} - \frac{1}{12}) V^{12}$</td>
<td>$-6321152A^9Z^2$</td>
<td>$-416608A^9Z^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-1787494A^8Z^3$</td>
<td>$-8339456A^8Z^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-3153600A^7Z^4$</td>
<td>$-10429440A^7Z^4$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-3831616A^6Z^5$</td>
<td>$-8866304A^6Z^5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-33582592A^5Z^6$</td>
<td>$-5214720A^5Z^6$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-21508096A^4Z^7$</td>
<td>$-2084864A^4Z^7$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-9965184A^3Z^8$</td>
<td>$-520576A^3Z^8$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-3211968A^2Z^9$</td>
<td>$-64832A^2Z^9$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-655424AZ^{10}$</td>
<td>$+64AZ^{10}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-65536Z^{11}$</td>
<td></td>
</tr>
</tbody>
</table>

Table E.2: The coefficients $I$ and $K$ of the generalized odd order Hamiltonian term in Eq. (6.10)
### Table E.3: The coefficients $R, S, T$ and $Q$ of the generalized even Hamiltonian term in Eqs. (6.55) and 6.40

<table>
<thead>
<tr>
<th>Order</th>
<th>Prefactor</th>
<th>$R$</th>
<th>$S$</th>
<th>$T$</th>
<th>$Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(1 - \frac{1}{2})V^2$</td>
<td>$-4A^2$</td>
<td>$-10AZ$</td>
<td>$+2AZ$</td>
<td>$+6Z^2$</td>
</tr>
<tr>
<td>2</td>
<td>$(\frac{1}{2} - \frac{1}{3})V^3$</td>
<td>$-8Z^2$</td>
<td>$+2AZ$</td>
<td>$-2Z^2$</td>
<td>$-88A^2Z^2$</td>
</tr>
<tr>
<td>3</td>
<td>$(\frac{1}{3} - \frac{1}{4})V^4$</td>
<td>$+16A^4$</td>
<td>$+152AZ$</td>
<td>$-88A^3Z$</td>
<td>$-88AZ^3$</td>
</tr>
<tr>
<td>4</td>
<td>$(\frac{1}{4} - \frac{1}{5})V^5$</td>
<td>$+272A^2Z^2$</td>
<td>$-88A^2Z^2$</td>
<td>$+44AZ^3$</td>
<td>$-52Z^4$</td>
</tr>
<tr>
<td>5</td>
<td>$(\frac{1}{5} - \frac{1}{6})V^6$</td>
<td>$-64A^6$</td>
<td>$-2208A^2Z$</td>
<td>$+1324A^3Z$</td>
<td>$+1632A^4Z^2$</td>
</tr>
<tr>
<td>6</td>
<td>$(\frac{1}{6} - \frac{1}{7})V^7$</td>
<td>$-6336A^4Z^2$</td>
<td>$-14320A^2Z^3$</td>
<td>$+3776A^2Z^3$</td>
<td>$+3264A^3Z^3$</td>
</tr>
<tr>
<td>7</td>
<td>$(\frac{1}{7} - \frac{1}{8})V^8$</td>
<td>$+256A^8$</td>
<td>$+33664A^7Z$</td>
<td>$-31616A^7Z$</td>
<td>$-29568A^6Z^2$</td>
</tr>
<tr>
<td>8</td>
<td>$(\frac{1}{8} - \frac{1}{9})V^9$</td>
<td>$+132608A^6Z^2$</td>
<td>$+95872A^5Z^2$</td>
<td>$-88704A^5Z^3$</td>
<td>$-13312A^4Z^4$</td>
</tr>
<tr>
<td>9</td>
<td>$(\frac{1}{9} - \frac{1}{10})V^{10}$</td>
<td>$+247360A^5Z^3$</td>
<td>$+247360A^5Z^3$</td>
<td>$-128384A^4Z^4$</td>
<td>$-118784A^3Z^5$</td>
</tr>
<tr>
<td>10</td>
<td>$(\frac{1}{10} - \frac{1}{11})V^{11}$</td>
<td>$+276616A^4Z^4$</td>
<td>$+276616A^4Z^4$</td>
<td>$+3952A^4Z^7$</td>
<td>$+2688A^2Z^7$</td>
</tr>
<tr>
<td>11</td>
<td>$(\frac{1}{11} - \frac{1}{12})V^{12}$</td>
<td>$+4096Z^8$</td>
<td>$+33664A^7Z$</td>
<td>$-31616A^7Z$</td>
<td>$-29568A^6Z^2$</td>
</tr>
</tbody>
</table>

The table lists the coefficients for orders 1 to 11 of the generalized even Hamiltonian term in Eqs. (6.55) and 6.40.
APPENDIX F

Comparison with the result of Kolley et al [55]

Kolley et al [55] used a Hamiltonian similar to the periodic Anderson model and carried out a canonical transformation similar to the Schrieffer and Wolff transformation to the fourth order, which is equivalent to our third order. Their objective was, however, slightly different to ours. In their paper, the canonical transformation was used to find out the higher order interactions of holes in the d and p orbitals in Cu-O lattices. Besides assuming strong on-site interaction, they also applied a single d-orbital occupation projection onto the final result, to keep the number of terms relatively minimal. One of the reasons for the large number of terms is that they kept all the indices independent, whereas we have assumed $i = i' = i''$ .... Secondly, they incorporated a phase factor into the hybridization term of the Hamiltonian, which introduces a minus sign when a p hole is created (or destroyed) on different side of an atom.

Nevertheless, there are still plenty of similarities between the two models. First of all, the p orbital holes in their model originate from oxygen ions which act like the free electrons in our model, whereas their d orbital holes from the Cu ions are like our impurities. Secondly, the on-site interaction, the energy of the d-orbitals and the energy of the p-orbitals, similar to that of the impurities and the free electrons in our Hamiltonian, are all included in their model. Furthermore, they assumed a two dimensional model which is compatible to our model.

To understand more about their superconducting version of the Anderson Model, the Hamiltonian is shown below in Eq. (F.1), and the $S$ required to satisfy Eq. (5.95) in Eq. (F.2). Since they used a different notation to ours, the symbols in the formulation below are different to the ones in their original paper. However the context is still the same. For example, the $t$ in Kolley's paper is equivalent to our $V$ and our $t$ is zero in their model.
F. Comparison with the result of Kolley et al [55]

\[
H = \epsilon_d \sum_{i \sigma} n_{i \sigma} + \epsilon_p \sum_{i \sigma} n_{i \sigma}^c + U \sum_i n_{i \sigma} f_{i \sigma}^d n_{i \sigma}^f + V \sum_{i \sigma} \sum_{l \neq i} (-1)^{\alpha_{il}} (f_{i \sigma}^c c_{l \sigma} + c_{l \sigma}^d f_{i \sigma})
\]

\[
S = V \sum_{i \sigma} \sum_{l \neq i} (-1)^{\alpha_{il}} \left[ \frac{1}{\epsilon} - \left( \frac{1}{\epsilon} + \frac{1}{U - \epsilon} \right) n_{i \sigma} f_{i \sigma}^d \right] (f_{i \sigma}^c c_{l \sigma} - c_{l \sigma}^d f_{i \sigma})
\]

The operator \( f_{i \sigma}^d (f_{i \sigma}) \), which is \( d_{i \sigma}^\dagger (d_{i \sigma}) \) in Kolley’s paper [55], creates (destroys) a hole with spin \( \sigma \) in the \( 3d_{x^2-y^2} \) orbital at the Cu site \( i \), whereas the \( c_{i \sigma}^\dagger (c_{i \sigma}) \), which is \( p_{i \sigma}^\dagger (p_{i \sigma}) \) in Kolley’s paper [55], creates (destroys) a hole with spin \( \sigma \) in the \( 2p_x \) or \( 2p_y \) orbital at the O site \( l \). \( \sum_{i (\neq i)} \) is the summation over the four nearest-neighbour sites around the Cu site \( i \) in a two-dimensional lattice involving one Cu and two O atoms per unit cell. \( \epsilon_d \) and \( \epsilon_p \) denote the atomic energies of d and p holes and \( U \) is the on-site interaction strength at Cu site. The Cu-O hybridization amplitude \( V \) has a phase factor associated with it, defined as

\[
(-1)^{\alpha_{il}} = \pm 1 \quad R_t = R_i = \frac{1}{2} e_z, \quad R_t = R_i = \frac{1}{2} e_y.
\]

This phase factor indeed comes from the radial wave solution of the Schrödinger equation which shows that holes or electrons at the top and left lobe of the atom have different sign to the ones at the bottom and right. Therefore if a p hole is created on the right of a destroyed d hole, a negative sign is needed to signify the sign change of the solution.

The result of their second order transformation is completely equivalent to ours when the phase factor and independent indices are incorporated into our formulation. However, going up to the third order with these two factors in our model is very difficult, if not impossible. We, therefore, took another approach to simplify their fourth order result with our assumption instead, before making a comparison. It turned out that only three out of six terms were identical. This is due to the single d-orbital occupation projection performed in their final result which caused the loss of terms. The two terms which were consistent in both results came from states that were derived without any double occupancy of d orbitals.

In the following, we show the comparison in detail by first writing their result in their second and fourth order and comparing them to our results term by term. To simply the coefficients, they adopted a new constant \( \epsilon \) to replace \( \epsilon_d \) and \( \epsilon_p \) as follow

\[
\epsilon = \epsilon_p - \epsilon_d
\]
which corresponds to our $A$ and $Z$ as

$$A = -\frac{1}{\varepsilon}, \quad Z = \frac{1}{U^2 - \varepsilon} + \frac{1}{\varepsilon} \quad \text{(F.5)}$$

Their second order result, equivalent to our first order, is

$$H'_{2} = H_{kin}^2 + H_{dp}^2 + H_{double}^2 \quad \text{(F.6)}$$
in which summing over only nearest neighbours $(i, j)$,

$$H_{kin}^2 = \frac{V^2}{\varepsilon} \left( -4 \sum_i n_i^{f} \right)$$

$$+ \sum_{i\sigma} \sum_{l(\neq i)} \left( \sum_{m(\neq i)} (-1)^{\alpha_{il} + \alpha_{lm}} c_{l\sigma}^{\dagger} c_{m\sigma} \right) \quad \text{(F.7)}$$

$$H_{dp}^2 = \frac{V^2}{\varepsilon} \left( \frac{1}{U - \varepsilon} \right) \sum_{i\sigma} \sum_{l(\neq i)} \left( \sum_{m(\neq i)} (-1)^{\alpha_{il} + \alpha_{lm}} \right)$$

$$\cdot \left( f_{i\sigma}^{f} f_{i\sigma}^{f} c_{l\sigma}^{\dagger} c_{m\sigma} - f_{i\sigma}^{f} f_{i\sigma}^{f} c_{l\sigma} c_{m\sigma} \right) \quad \text{(F.8)}$$

$$H_{double}^2 = \frac{V^2}{\varepsilon} \sum_{(i,j)\sigma} f_{i\sigma}^{f} f_{j\sigma}^{f} + \frac{V^2}{2} \left( \frac{1}{U - \varepsilon} \right)$$

$$\cdot \left( 8 \sum_{i\sigma} n_{i\sigma}^{f} n_{i\sigma}^{f} - \sum_{<i,j>\sigma} f_{i\sigma}^{f} f_{j\sigma}^{f} (n_{i\sigma}^{f} + n_{j\sigma}^{f}) \right)$$

$$+ \sum_{i\sigma} \sum_{l(\neq i)} \left( \sum_{m(\neq i)} (-1)^{\alpha_{il} + \alpha_{lm}} \right)$$

$$\cdot \left( f_{i\sigma}^{f} f_{i\sigma}^{f} c_{m\sigma} + c_{l\sigma}^{\dagger} c_{m\sigma} f_{i\sigma}^{f} f_{i\sigma}^{f} d_{i\sigma} \right) \quad \text{(F.9)}$$

The equivalence of our result to theirs is apparent if we substitute $\frac{1}{\varepsilon}$ with $-A$ and $\frac{1}{U^2 - \varepsilon} + \frac{1}{\varepsilon}$ with $Z$, and adding the phase factor in the $H$ and $S$ of our model as in Eqs. (F.1) and (F.2)\footnote{When re-deriving $H_V = -[S, H_0]$, it is interesting to note that the phase factor can always be taken out of the summation. But in the next order transformation, the phase factors with different indices are different in different terms and hence the summation which could be easily done in the first order, becomes next to impossible.}. With the fourth order result, our assumption on the indices being equal is first applied, with the $(-1)^{\alpha_{ij}}$ factor ignored. This step is quite tricky since Kolleys et al wrote the summation of $i, j$ as four times the summation of $i$ with $j = i$ and the summation of $(i, j)$ with $i$ and $j$ different. In order to assume $i$ and $j$ being equal, we have to take the first sum and divide it by four, with the second sum completely neglected. However if there is only one index $i$ and some indices other than $j$ in the
summation, no alternation is needed, except for the term $n^f_{i\sigma}$. Since the sum of this term has only one index, Kolley et al would have to multiply the 4 by 5 to take into account the nearest neighbour and the double counting. After this modification, their result becomes

$$H^4 = H^4_{\text{kin}} + H^4_{dp} + H^4_{dd} + H^4_{pp} + H^4_{ddp}$$  \hspace{1cm} (F.10)

where

$$H^4_{\text{kin}} = \frac{V^3}{\epsilon^3} \sum_{i,\sigma} (n^f_{i\sigma} - n^c_{i\sigma})$$  \hspace{1cm} (F.11)

$$H^4_{dp} = -\frac{1}{4} J_{dp} \sum_{i\sigma} (f^\dagger_{i,\sigma} f_{i,-\sigma} c^\dagger_{i,-\sigma} c_{i,\sigma} - f^\dagger_{i,-\sigma} f_{i,\sigma} c^\dagger_{i,\sigma} c_{i,-\sigma})$$  \hspace{1cm} (F.12)

$$H^4_{dd} = \frac{1}{4} J_{dd} \sum_{i\sigma} (f^\dagger_{i,\sigma} f_{i,-\sigma} f^\dagger_{i,-\sigma} f_{i,\sigma} - f^\dagger_{i,\sigma} f_{i,\sigma} f^\dagger_{i,-\sigma} f_{i,-\sigma})$$  \hspace{1cm} (F.13)

$$H^4_{pp} = -\sum_{i\sigma} (J_{pp} - \frac{1}{2} J_{ddp} n^d_{i\sigma} c^\dagger_{i,\sigma} c^\dagger_{i,-\sigma} c_{i,-\sigma} c_{i,\sigma})$$  \hspace{1cm} (F.14)

$$H^4_{ddp} = J_{ddp} \sum_{i\sigma} (f^\dagger_{i,-\sigma} f_{i,\sigma} f^\dagger_{i,-\sigma} f_{i,\sigma} - f^\dagger_{i,-\sigma} f_{i,-\sigma} f^\dagger_{i,-\sigma} f_{i,-\sigma} c^\dagger_{i,\sigma} c^\dagger_{i,-\sigma} c_{i,-\sigma} c_{i,\sigma})$$  \hspace{1cm} (F.15)

$$J_{dp} = 8V^4 \left( \frac{1}{\epsilon^3} + \frac{1}{2 \epsilon^2 (U - \epsilon)} - \frac{1}{2 \epsilon (U - \epsilon)^2} + \frac{1}{(U - \epsilon)^3} \right)$$  \hspace{1cm} (F.16)

$$J_{dd} = V^4 \left( \frac{2}{\epsilon^3} + \frac{3}{2 \epsilon^2 (U - \epsilon)} - \frac{1}{2 \epsilon (U - \epsilon)^2} \right)$$  \hspace{1cm} (F.17)

$$J_{pp} = V^4 \left( \frac{2}{\epsilon^3} + \frac{3}{2 \epsilon^2 (U - \epsilon)} + \frac{1}{\epsilon (U - \epsilon)^2} \right)$$  \hspace{1cm} (F.18)

$$J_{ddp} = V^4 \left( \frac{1}{\epsilon^3} + \frac{1}{\epsilon^2 (U - \epsilon)} - \frac{1}{\epsilon (U - \epsilon)^2} - \frac{1}{(U - \epsilon)^3} \right)$$  \hspace{1cm} (F.19)

Substituting $A = -\frac{1}{\epsilon}$ and $A + Z = \frac{1}{U - \epsilon}$, we get the coefficients of the operators $f^\dagger_{i,\sigma} f_{i,-\sigma} c^\dagger_{i,-\sigma} c_{i,\sigma} - f^\dagger_{i,-\sigma} f_{i,\sigma} c^\dagger_{i,\sigma} c_{i,-\sigma}$:

$$-\frac{1}{4} J_{dp} = \frac{1}{8} V^4 (-32A^2Z - 32AZ^2 - 16Z^3) = J_{m=1}$$  \hspace{1cm} (F.21)

the coefficient of the operators $n^c_{i\sigma} n^c_{i,-\sigma}$

$$-J_{pp} = \frac{1}{8} V^4 (-4A^2Z + 4A^2Z) = K_{m=1}$$  \hspace{1cm} (F.22)

and the coefficient of the operator $n^f_{i\sigma} - n_{i\sigma}$:

$$\frac{1}{8} (8V^4 A^3 = G_{m=1}$$  \hspace{1cm} (F.23)
APPENDIX G

The extended Anderson lattice model

This appendix presents the results of an extended Anderson lattice model subject to a canonical transformation. This model contains terms from the standard Anderson model, as well as an extra term with on-site free electron interaction.

\[ H = H_0 + H_V \] (G.1)

where

\[ H_0 = \sum_{\langle i \sigma \rangle} \left( t(c_{i+1\sigma}^\dagger c_{i\sigma} + c_{i\sigma}^\dagger c_{i+1\sigma}) - \mu c_{i\sigma}^\dagger c_{i\sigma} + \epsilon_f f_{i\sigma}^\dagger f_{i\sigma} + U f_{i\uparrow}^\dagger f_{i\uparrow} f_{i\downarrow}^\dagger f_{i\downarrow} + W c_{i\uparrow}^\dagger c_{i\downarrow} c_{i\downarrow}^\dagger c_{i\uparrow} \right) \] (G.2)

\[ H_V = \sum_{\langle \sigma \rangle} V (f_{i\sigma}^\dagger c_{i\sigma} + c_{i\sigma}^\dagger f_{i\sigma}) \] (G.3)

The on-site Coulomb interaction between free electrons is represented by the new term \( W c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger c_{i\downarrow} c_{i\uparrow} \) which increases the energy of the system with double occupation of free electrons at the same site, provided the positivity of \( W \). When \( W \) is set to zero, this model is identical to the periodic Anderson model used for the rest of the paper. Therefore, this model can be considered as a more generic version of the periodic Anderson model.

The canonical transformation is the same as before

\[ \tilde{H} = e^S H e^{-S} \]

\[ = H_0 + H_V + [S, H_0] + [S, H_V] + \frac{1}{2}[S, [S, H_0]] + \frac{1}{2}[S, [S, H_V]] + \cdots \] (G.4)

to which \( S \) needs to satisfy

\[ [S, H_0] = -H_V \] (G.5)
It can again be written in the form

\[
S = \sum_{i',j',\sigma'} \left( A + Z n_{j',-\sigma'}^f + X n_{j',-\sigma'}^e + Y n_{j',-\sigma'}^f \right) \left( V_{i',j',\sigma'} c_{i',\sigma'}^e - V_{i',j',\sigma'} c_{i',\sigma'}^f \right),
\]

(G.6)

where \( A, Z, X \) and \( Y \) are determined from the condition in Eq. (G.5). The derivation of the coefficients from the canonical transformation is very similar to that of chapter 6, and hence will be skipped here. What is important here are the differences of the coefficients from that of the standard periodic Anderson model. The new results from the new formulation of \( S \) will therefore be presented here without detail derivation.

Again the Hamiltonian will be divided into two parts, the odd and the even.

\[
\hat{H} = H_0 + H_{\text{odd}} + H_{\text{even}},
\]

(G.7)

where

\[
H_{\text{odd}} = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] ([S, H_V])_{2m+1},
\]

(G.8)

\[
H_{\text{even}} = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+2)!} - \frac{1}{(2m+3)!} \right] ([S, H_V])_{2m+2}.
\]

(G.9)

The even and odd Hamiltonians can be simplified as

\[
H_{\text{odd}} = \sum_{i,\sigma} \left( J''(c_{i,\sigma} c_{i,-\sigma}^f f_{i,\sigma} - n_{i,\sigma}^f n_{i,-\sigma}^f) + P''(c_{i,\sigma} c_{i,-\sigma}^f f_{i,\sigma} f_{i,-\sigma}) + f_{i,\sigma}^f f_{i,-\sigma}^f c_{i,\sigma} c_{i,-\sigma} + G''(n_{i,\sigma}^f - n_{i,-\sigma}^e) + I''(n_{i,-\sigma}^f n_{i,\sigma}^f + K'' n_{i,-\sigma}^e n_{i,\sigma}^e) + M''(n_{i,\sigma}^f n_{i,-\sigma}^e n_{i,-\sigma}^f - n_{i,\sigma}^e) \right),
\]

(G.10)

where \( J'', P'', G'', I'', K'', M'' \) are the summation of the corresponding \( J''_m, P''_m, G''_m, I''_m, K''_m, M''_m \) over infinite number of \( m \)'s. If we define \( \theta = \sqrt{2V^2(\mu^2 + \nu^2)}, \theta_\sigma = 2V\gamma \) and \( \theta_\xi = 2V\xi \),

\[
J'' = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] \cdot 2(\mu - \nu)V^2(-2^3V^2(\mu^2 + \nu^2))^{m} = 2(\mu - \nu)V^2 \left( \frac{\sin 2\theta}{2\theta} + \frac{\cos 2\theta - 1}{(2\theta)^2} \right).
\]

(G.11)
\begin{align*}
P'' &= \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] \left( 2\mu\nu \frac{\mu + \nu}{\mu^2 + \nu^2} V^2 (-2^3 V^2 (\mu^2 + \nu^2))^m ight. \\
&\left. \quad - (\mu - \nu V^2 (\mu + \nu)^2 \mu^2 + \nu^2 (-2V^2 (\mu^2 + \nu^2))^m \right) \\
&= 2\mu\nu (\mu - \nu V^2 (\mu + \nu)^2 \mu^2 + \nu^2 (\sin \frac{2\theta}{2\theta} + \cos \frac{2\theta - 1}{(\theta)^2}) \\
&\left. \quad - (\mu - \nu V^2 (\mu + \nu)^2 \mu^2 + \nu^2 \frac{\sin \theta}{\theta} + \cos \theta - \frac{1}{(\theta)^2}) \right) \\
&= 2\gamma V^2 (\sin \frac{\theta}{\theta} + \cos \theta - \frac{1}{(\theta)^2}) \
\end{align*}

\begin{align*}
G'' &= \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] 2\gamma V^2 (-2V^2 (2\gamma^2))^m \\
&= 2\gamma V^2 (\sin \frac{\theta}{\theta} + \cos \theta - \frac{1}{(\theta)^2}) \
\end{align*}

\begin{align*}
K'' &= \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] \left( 2\nu^2 V^2 \frac{\mu - \nu}{\mu^2 + \nu^2} (-2^3 V^2 (\mu^2 + \nu^2))^m ight. \\
&\left. \quad - 2\mu\nu \frac{\mu + \nu}{\mu^2 + \nu^2} V^2 (-2V^2 (\mu^2 + \nu^2))^m + 2\gamma V^2 (-2V^2 (2\gamma^2))^m \right) \\
&= 2\nu^2 V^2 \frac{\mu - \nu}{\mu^2 + \nu^2} \left( \frac{\sin \frac{2\theta}{2\theta} + \cos \frac{2\theta - 1}{(\theta)^2}}{(\theta)^2} \right) \\
&\left. \quad - 2\mu\nu V^2 \frac{\mu + \nu}{\mu^2 + \nu^2} \left( \frac{\sin \frac{\theta}{\theta} + \cos \frac{\theta - 1}{(\theta)^2}}{(\theta)^2} \right) + G \right) \\
&= 2\mu^2 V^2 \frac{\mu - \nu}{\mu^2 + \nu^2} \left( \frac{\sin \frac{2\theta}{2\theta} + \cos \frac{2\theta - 1}{(\theta)^2}}{(\theta)^2} \right) \\
&\left. \quad + 2\mu\nu V^2 \frac{\mu + \nu}{\mu^2 + \nu^2} \left( \frac{\sin \frac{\theta}{\theta} + \cos \frac{\theta - 1}{(\theta)^2}}{(\theta)^2} \right) \right) - G \\
\end{align*}

\begin{align*}
I'' &= \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] \left( 2\mu^2 V^2 \frac{\mu - \nu}{\mu^2 + \nu^2} (-2^3 V^2 (\mu^2 + \nu^2))^m ight. \\
&\left. \quad + 2\mu\nu V^2 \frac{\mu + \nu}{\mu^2 + \nu^2} (-2V^2 (\mu^2 + \nu^2))^m + 2\gamma V^2 (-2V^2 (2\gamma^2))^m \right) \\
&= 2\mu^2 V^2 \frac{\mu - \nu}{\mu^2 + \nu^2} \left( \frac{\sin \frac{2\theta}{2\theta} + \cos \frac{2\theta - 1}{(\theta)^2}}{(\theta)^2} \right) \\
&\left. \quad + 2\mu\nu V^2 \frac{\mu + \nu}{\mu^2 + \nu^2} \left( \frac{\sin \frac{\theta}{\theta} + \cos \frac{\theta - 1}{(\theta)^2}}{(\theta)^2} \right) \right) - G \\
\end{align*}
\[ M'' = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+1)!} - \frac{1}{(2m+2)!} \right] \left( -2(\mu - \nu) \frac{\mu^2 - \nu^2}{\mu^2 + \nu^2} V^2 (-2^3 V^2)^m \right) \]
\[
\cdot (\mu^2 + \nu^2)^m - 4\mu\nu \frac{\mu + \nu}{\mu^2 + \nu^2} V^2 (-2V^2)^m (\mu^2 + \nu^2)^m \\
+ 2\xi V^2 (-4\xi^2 V^2)^m + 2\gamma V^2 (-4\gamma^2 V^2)^m \]
\[
= -2(\mu - \nu) \frac{\mu^2 - \nu^2}{\mu^2 + \nu^2} V^2 \left( \frac{\sin 2\theta}{2\theta} + \frac{\cos 2\theta - 1}{(2\theta)^2} \right) - 4\mu\nu \frac{\mu + \nu}{\mu^2 + \nu^2} V^2 \\
\cdot \left( \frac{\sin \theta}{\theta} + \frac{\cos \theta - 1}{\theta^2} \right) + 2\xi V^2 \left( \frac{\sin \xi}{\xi} + \frac{\cos \xi - 1}{\xi^2} \right) \\
+ 2\gamma V^2 \left( \frac{\sin \gamma}{\gamma} + \frac{\cos \gamma - 1}{\gamma^2} \right) \quad \text{(G.16)}
\]

and

\[ H_{even} = \sum_{i\sigma} (R'' + S''n_{i-\sigma}^f + T''n_{i-\sigma}^c + W''n_{i-\sigma}^f n_{i-\sigma}^c) (c_{i\sigma}^f f_{i\sigma} + f_{i\sigma}^\dagger c_{i\sigma}) \quad \text{(G.17)} \]

where \( R'' \), \( S'' \), \( W'' \), \( T'' \) are the summation of the corresponding \( R_m'', S_m'', Q_m'', T_m'' \) over infinite number of \( m \)'s. If we use the same \( \theta, \theta_\gamma, \theta_\xi \) as in the odd order coefficients,

\[ R'' = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+2)!} - \frac{1}{(2m+3)!} \right] (-4\gamma^2 V^3 (-4V^2\gamma^2)^m) \]  
\[ = -4\gamma^2 V^3 \left( \frac{\sin \theta_\gamma}{\theta_\gamma^3} - \frac{\cos \theta_\gamma}{\theta_\gamma^2} \right) \quad \text{(G.18)} \]

\[ S'' = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+2)!} - \frac{1}{(2m+3)!} \right] \left( -8\mu(\mu - \nu)V^3 (-2^3 V^2(\mu^2 + \nu^2))^m \right) \]
\[
-2\mu(\mu + \nu)V^3 (-2V^2(\mu^2 + \nu^2))^m + 4\gamma^2 V^3 (-4V^2\gamma^2)^m \]
\[
= -8\mu(\mu - \nu)V^3 \left( \frac{\sin 2\theta}{(2\theta)^3} - \frac{\cos 2\theta}{(2\theta)^2} \right) \\
-2\mu(\mu + \nu)V^3 \left( \frac{\sin \theta}{\theta^3} - \frac{\cos \theta}{\theta^2} \right) - R \quad \text{(G.19)}
\]
\[ T'' = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+2)!} - \frac{1}{(2m+3)!} \right] \left( 8\nu(\mu - \nu)V^3(-2^3V^2(\mu^2 + \nu^2))^m 
- 2\mu(\mu + \nu)V^3(-2V^2(\mu^2 + \nu^2))^m + 4\gamma^2V^3(-4V^2\gamma^2)^m \right) \]

\[ = 8\nu(\mu - \nu)V^3\left( \frac{\sin 2\theta}{(2\theta)^3} - \frac{\cos 2\theta}{(2\theta)^2} \right) \]
\[ - 2\mu(\mu + \nu)V^3\left( \frac{\sin \theta}{\theta^3} - \frac{\cos \theta}{\theta^2} \right) - R \]

\[ (G.20) \]

\[ Q'' = \sum_{m=0}^{\infty} \left[ \frac{1}{(2m+2)!} - \frac{1}{(2m+3)!} \right] \left( 8(\mu - \nu)^2V^3(-2^3V^2(\mu^2 + \nu^2))^m 
+ 2(\mu + \nu)^2V^3(-2V^2(\mu^2 + \nu^2))^m - 4\xi^2V^3(-4\xi^2V^2)^m - 4\gamma^2V^3 \right) \]

\[ = 8(\mu - \nu)^2V^3\left( \frac{\sin 2\theta}{(2\theta)^3} - \frac{\cos 2\theta}{(2\theta)^2} \right) + 2(\mu + \nu)^2V^3\left( \frac{\sin \theta}{\theta^3} - \frac{\cos \theta}{\theta^2} \right) \]
\[ - 4\gamma V^3\left( \frac{\sin \theta_\gamma}{\theta_\gamma^3} - \frac{\cos \theta_\gamma}{\theta_\gamma^2} \right) - 4\xi V^3\left( \frac{\sin \theta_\xi}{\theta_\xi^3} - \frac{\cos \theta_\xi}{\theta_\xi^2} \right) \]

\[ (G.21) \]

These results are derived based upon the condition in Eq. (G.5) which however only eliminates the first order hybridization out of the Hamiltonian. If we now follow the same approach as in chapter 7 and assume the parameters \(A, X, Y\) and \(Z\) of the most generic form, we can solve for them from four different conditions that would set all hybridization terms to zero, namely \(R = S = T = Q = 0\).

We begin with the first commutation

\[ [S, H_0] = \sum_j V(A' + Z'n_{j-\sigma}^i + X'n_{j-\sigma}^e + Y'n_{j-\sigma}^i)(f_j^\dagger c_j + c_j^\dagger f_j) \]

in which

\[ A' = (2t - \mu_c - \epsilon_f)A, \]
\[ Z' = (2t - \mu_c - \epsilon_f)Z - UA - UZ, \]
\[ X' = (2t - \mu_c - \epsilon_f)X + WA + WZ, \]
\[ Y' = (2t - \mu_c - \epsilon_f)Y - UY - UZ + WZ + WY. \]

The result is very similar to the even order transformation pattern and falls into the same format as outlined in chapter 6. We therefore use the same procedure to
The extended Anderson lattice model determine the next order commutation.

\[
[S,[S,H_0]] = \sum_{i, \sigma} \left[ J_0'(c_{i, \sigma}^\dagger c_{i,-\sigma} f_{i,-\sigma}^\dagger f_{i,\sigma} - n_{i, \sigma}^c n_{i,-\sigma}^c) + G_0'(n_{i, \sigma}^f - n_{i,-\sigma}^f) + P_0'(c_{i, \sigma}^\dagger c_{i,-\sigma} f_{i,-\sigma}^\dagger f_{i,\sigma} + f_{i, \sigma}^\dagger f_{i,-\sigma} c_{i, \sigma} c_{i,-\sigma}) + I_0' n_{i,-\sigma}^f n_{i, \sigma}^f \\
+ M_0' n_{i,-\sigma}^c n_{i,-\sigma}^c (n_{i, \sigma}^f - n_{i,-\sigma}^f) + K_0' n_{i,-\sigma}^c n_{i, \sigma}^c \right],
\]

which, using a short hand notation of \( \mu = A + Z, \nu = A + X, \gamma = A, \rho = A + Z + X + Y \) and \( \xi = 2t - \mu c - \epsilon_f \) gives

\[
J_0' = 2V^2(\xi(\mu^2 - \nu^2) - U\mu^2 - W\nu^2); \\
P_0' = -V^2(W + U)\mu\nu; \\
G_0' = 2V^2\gamma^2\xi; \\
I_0' = 2V^2(\xi\mu^2 - \xi\gamma^2 - U\mu^2); \\
K_0' = -2V^2(\xi\nu^2 - \xi\gamma^2 + W\nu^2); \\
M_0' = 2V^2(\xi(\rho^2 - \mu^2 - \nu^2 + \gamma^2) - U(\rho^2 - \mu^2) + W(\rho^2 - \nu^2)).
\]

These results are very similar to the results of the periodic Anderson model, except the initial conditions and the value of different parameters. The Hamiltonian after the canonical transformation can therefore be evaluated using the same algebra as in chapter 7, in which

\[
\hat{H} = H_0 + H_V + [S,H_0] + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} [S,[S,H_0]_n] + \sum_{n=1}^{\infty} \frac{1}{(n)!} [S,H_V]_n
\]

This equation can also be written as the sum of an odd and an even part defined as

\[
H_{odd} = \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} [S,H_V]_{2m+1} + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} [S,[S,H_0]_{2m+1}
\]

\[
H_{even} = H_V + [S,H_0] + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} [S,H_V]_{2m+2}
\]

and then summed to all orders, similar to the periodic Anderson model, to yield Eqs. (6.48) and (6.55) with different values of \( J,P,G,M,I,K \) for odd orders and \( R,S,T,Q \) for even orders. For instance, the coefficient \( R \) in \( H_{even} \) will be identical to that of the periodic Anderson model (7.53), as all parameters are identical with respect to \( R \).

\[
R^{(3)} = V + A'V + (\cos(2VA) - 1)V + (\sin(2VA)/2VA - 1)VA'
\]
With $S$, the calculation becomes more complicated in the summation,

\[ S^{(3)} = VZ' + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} S_m + \sum_{m=0}^{\infty} \frac{1}{(2m+3)!} S'_m \]  

(G.38)

which finally yields after substitution

\[ S^{(3)} = 8V^3 \mu(\mu^2 - \nu^2) - U \mu^2 - W \nu^2) \phi(\sqrt{8V^2(\mu^2 + \nu^2)}) + 8V^3 \mu(\mu - \nu) \psi(\sqrt{8V^2(\mu^2 + \nu^2)}) + 2V^3 \mu \nu^2 (2\xi - U + W) \phi(\sqrt{2V^2(\mu^2 + \nu^2)}) + 2V^3 \nu(\mu + \nu) \psi(\sqrt{2V^2(\mu^2 + \nu^2)}) - V \cos 2V\gamma - \frac{\xi}{2} \sin 2V\gamma \]  

(G.39)

where different to defined before $\phi(x) = \sin(x)/x^3$ and $\psi(x) = \cos(x)/x^2$.

Similarly for $T$,

\[ T^{(3)} = VX' + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} T_m + \sum_{m=0}^{\infty} \frac{1}{(2m+3)!} T'_m \]  

(G.40)

yields

\[ T^{(3)} = -8V^3 \nu(\xi(\mu^2 - \nu^2) - U \mu^2 - W \nu^2) \phi(\sqrt{8V^2(\mu^2 + \nu^2)}) - 8V^3 \nu(\mu - \nu) \psi(\sqrt{8V^2(\mu^2 + \nu^2)}) + 2V^3 \mu \nu^2 (2\xi - U + W) \phi(\sqrt{2V^2(\mu^2 + \nu^2)}) + 2V^3 \nu(\mu + \nu) \psi(\sqrt{2V^2(\mu^2 + \nu^2)}) - V \cos 2V\gamma - \frac{\xi}{2} \sin 2V\gamma. \]  

(G.41)

Finally, the value $Q$ is given by

\[ Q^{(3)} = VY' + \sum_{m=0}^{\infty} \frac{1}{(2m+2)!} Q_m + \sum_{m=0}^{\infty} \frac{1}{(2m+3)!} Q'_m \]  

(G.42)

which after substitution, yields

\[ Q^{(3)} = -8V^3(\mu - \nu)(\xi(\mu^2 - \nu^2) - U \mu^2 - W \nu^2) \phi(\sqrt{8V^2(\mu^2 + \nu^2)}) - 8V^3(\mu - \nu^2) \psi(\sqrt{8V^2(\mu^2 + \nu^2)}) - 2V^3 \mu(\mu + \nu)(2\xi - U + W) \phi(\sqrt{2V^2(\mu^2 + \nu^2)}) - 2V^3(\nu(\mu + \nu) \psi(\sqrt{2V^2(\mu^2 + \nu^2)}) - V(\gamma^2 + \rho^2) \cos 2V\rho + \frac{1}{2}(\xi(\mu^2 + \nu^2) + (\xi - U + W)) \sin 2V\rho + V(1 + \xi(\gamma^2 - \rho^2)) \]  

(G.43)
APPENDIX H

The transformation results in asymmetric case

This appendix contains plots of the coefficients in the asymmetric case when \( \epsilon_f \) is no longer equal to \( U/2 \). This condition increases the number of free variables from two to three, namely \( U, V \) and \( \epsilon_f \), and hence three plots are needed for each coefficient to show the three degrees of freedom. The parameters other than \( U, V \) and \( \epsilon_f \) are fixed to zero, i.e. \( t = \mu = 0 \). The solid lines in the plots below show the coefficients of the Hamiltonian after canonical transformation when \( R = S \) are set to zero. The long solid dashed lines in the plots show the coefficients after transformation when \( R, S, T \) and \( Q \) are all set to zero. The grey long dashed lines show the coefficients after transformation with only the first order hybridization removed. The grey short dashed lines show the coefficients of the first order Schrieffer Wolff transformation result.

In figure H.1, the first order singularity in the Schrieffer Wolff result and the high order singularity in our infinite sum result at \( U = 3 \) have been removed from the two transformations, one set \( R \) and \( S \) (solid line) to zero and the other all \( R, S, T \)

Figure H.1: Coefficient \( J \) vs \( U \), Figure H.2: Coefficient \( J \) vs \( V \), Figure H.3: Coefficient \( J \) vs \( \epsilon_f \), with \( t = 0, \mu = 0, \epsilon_F = -3 \) and \( \epsilon_f = \) with \( t = 0, \mu = 0, V = 1 \) and \( V = 1 \). See context for details. \( U = 8 \). See context for details. \( U = 8 \). See context for details.
Figure H.4: Coefficient $P$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3$ $V$, with $t = 0, \mu = 0, \epsilon_F = -3$ $e_f$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

Figure H.5: Coefficient $P$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3$ $V$, with $t = 0, \mu = 0, \epsilon_F = -3$ $e_f$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

Figure H.6: Coefficient $P$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3$ $V$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

and $Q$ (solid long dashed) to zero. The same property can also be seen in figure H.4, H.10, H.13, H.16, H.22, H.25 and figure H.28.

The plots of the coefficients as a function of $V$ show the oscillatory behaviour away from the singular point in grey dashed lines. The solid and the solid long dashed lines both show the absence of oscillation after the removal of the singularity from the transformations which set $R$ and $S$ (the solid line) to zero and all $R, S, T$ and $Q$ (the solid long dashed) to zero. In two cases with the coefficients $P$ in figure H.5 and $G$ in figure H.8, the solid line and the solid dashed line coincide with each other which imply the setting of coefficients $T$ and $Q$ to zero has no effect to these two coefficients.

The coefficients as a function of $\epsilon_f$ show two singular points at $\epsilon_f = 0$ and $-8$ in most cases, except $G$ in figure H.9, $K$ in figure H.15 and $R$ in figure H.21, for the Schrieffer Wolff transformation and our infinite sum version of the Schrieffer Wolff transformation. The two singularities are due to the two poles in the parameter $S$.

Figure H.7: Coefficient $G$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3$ $V$, with $t = 0, \mu = 0, \epsilon_F = -3$ $e_f$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

Figure H.8: Coefficient $G$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3$ $V$, with $t = 0, \mu = 0, \epsilon_F = -3$ $e_f$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

Figure H.9: Coefficient $G$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3$ $V$, with $t = 0, \mu = 0, \epsilon_F = -3$ $e_f$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.
one in the variable $A$ and the other in the variable $Z$. The removal of one of the two singularities in the above mentioned three cases is because of the single dependence of the corresponding coefficient to one variable $A$ only. It is also obvious to see that after using the transformation which set either $R$ and $S$ (the solid line) to zero or all $R, S, T$ and $Q$ (the solid long dashed) to zero, one of the two singularities has turn into a phase shift of $\pi$ while the other one has disappeared. The phase shift is in fact an artefact of the arctan function. One can simply remove the phase shift by adding or subtracting $\pi$ to the rest of the curve to make it continuous as arctan is periodic with a cycle of $\pi$.

Finally the solid dashed lines are always sitting at zero for the coefficients $R, S, T$ and $Q$ from figure H.19 to figure H.30 by definition, but only the solid lines are sitting at zero for the coefficients $R$ and $S$ from figure H.19 to figure H.24. The fact that the solid lines of the coefficients $T$ and $Q$ are non-zero from figure H.25 to H.30 indicates that the hybridization terms still survive through the transformation
which only sets the coefficients $R$ and $S$ to zero.

Figure H.16: Coefficient $M$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3\,\text{V}$, with $t = 0, \mu = 0, \epsilon_F = -3\,\text{e}_f$, with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

Figure H.17: Coefficient $M$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3\,\text{V}$, with $t = 0, \mu = 0, \epsilon_F = -3\,\text{e}_f$, with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.

Figure H.18: Coefficient $M$ vs $U$, with $t = 0, \mu = 0, \epsilon_F = -3\,\text{V}$, with $t = 0, \mu = 0, \epsilon_F = -3\,\text{e}_f$, with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.

Figure H.19: Coefficient $R$ vs $U$ with $t = 0, \mu = 0, \epsilon_F = -3\,\text{V}$ with $t = 0, \mu = 0, \epsilon_F = -3\,\text{e}_f$ with $t = 0, \mu = 0, V = 1$ and $V = 1$. See context for details.

Figure H.20: Coefficient $R$ vs $U$ with $t = 0, \mu = 0, \epsilon_F = -3\,\text{V}$ with $t = 0, \mu = 0, \epsilon_F = -3\,\text{e}_f$ with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.

Figure H.21: Coefficient $R$ vs $U$ with $t = 0, \mu = 0, \epsilon_F = -3\,\text{V}$ with $t = 0, \mu = 0, \epsilon_F = -3\,\text{e}_f$ with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.
Figure H.22: Coefficient $S$ vs $U$ with $t = 0, \mu = 0, \epsilon_F = -3$ and $V = 1$. See context for details.

Figure H.23: Coefficient $S$ vs $V$ with $t = 0, \mu = 0, \epsilon_F = -3$ and $U = 8$. See context for details.

Figure H.24: Coefficient $S$ vs $\eta$ with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.

Figure H.25: Coefficient $T$ vs $U$ with $t = 0, \mu = 0, \epsilon_F = -3$ and $V = 1$. See context for details.

Figure H.26: Coefficient $T$ vs $V$ with $t = 0, \mu = 0, \epsilon_F = -3$ and $U = 8$. See context for details.

Figure H.27: Coefficient $T$ vs $\eta$ with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.

Figure H.28: Coefficient $Q$ vs $U$ with $t = 0, \mu = 0, \epsilon_F = -3$ and $V = 1$. See context for details.

Figure H.29: Coefficient $Q$ vs $V$ with $t = 0, \mu = 0, \epsilon_F = -3$ and $U = 8$. See context for details.

Figure H.30: Coefficient $Q$ vs $\eta$ with $t = 0, \mu = 0, V = 1$ and $U = 8$. See context for details.
The fourth and fifth order results of non-crossing hopping

Following the calculation of canonical transformation for non-crossing hopping from chapter 7, one can in general find out the canonical transformation of the periodic Anderson model to any order, assuming the particles created from each order do not collide with one another. The first three order results can easily be done by hand, but the fourth and fifth are more difficult. We will follow the steps described in chapter 7 to work out the fourth and fifth order in this appendix.

For the fifth order result, the corresponding matrix \( M \) in Eq. (7.15) is given by

\[
M = \begin{pmatrix}
1 & \alpha^\sigma & 4Z\alpha^\sigma & 6Z^2\alpha^\sigma & 4Z^3\alpha^\sigma & Z^4\alpha^\sigma \\
0 & 1 & \alpha^\sigma & 3Z\alpha^\sigma & 3Z^2\alpha^\sigma & Z^3\alpha^\sigma \\
0 & 0 & 1 & \alpha^\sigma & 2Z\alpha^\sigma & Z^2\alpha^\sigma \\
0 & 0 & 0 & 1 & \alpha^\sigma & Z\alpha^\sigma \\
0 & 0 & 0 & 0 & 1 & \alpha^\sigma \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

and the solution to the equation \( M \cdot B = C \) in Eq.(7.15) can be found from the co-factors of the elements from the last column of the matrices \( M_1, M_2, \ldots \) All vectors and matrixes are the same as the ones defined in chapter 7. For instance, to find the solution of \( \hat{\phi}^5\alpha^\sigma \) one needs to find the determinant of \( M_1 \), the same matrix as \( M \) except that the first column is replaced by the column vector \( C \). The evaluation of this determinant is carried out using the cofactors of the elements from
the last column. In other words, the determinant of $M_1$ is given by

\[
-\beta^{-\sigma} Z \hat{O}_{4} \alpha^\sigma
\]

\[
\begin{vmatrix}
1 & Z \alpha^\sigma & 3Z \hat{O}_{\alpha^\sigma} & 3Z \hat{O}_{2} \alpha^\sigma \\
0 & 1 & Z \alpha^\sigma & 2Z \hat{O}_{\alpha^\sigma} \\
0 & 0 & 1 & Z \alpha^\sigma \\
0 & 0 & 0 & 1 \\
\end{vmatrix}
\]

\[
+\beta^{-\sigma} Z \hat{O}_{3} \alpha^\sigma
\]

\[
\begin{vmatrix}
Z \alpha^\sigma & 4Z \hat{O}_{\alpha^\sigma} & 6Z \hat{O}_{2} \alpha^\sigma & 4Z \hat{O}_{3} \alpha^\sigma \\
0 & 1 & Z \alpha^\sigma & 2Z \hat{O}_{\alpha^\sigma} \\
0 & 0 & 1 & Z \alpha^\sigma \\
0 & 0 & 0 & 1 \\
\end{vmatrix}
\]

The determinant for the first term of this expression is $-\beta^{-\sigma} Z \hat{O}_{4} \alpha^\sigma$, and the determinant for the second term is $\beta^{-\sigma} Z^2 \alpha^\sigma \hat{O}_{3} \alpha^\sigma$. Following the same logic, we can write

\[
|M_1| = -\beta^{-\sigma} (Z \hat{O}_{4} \alpha^\sigma - Z \hat{O}_{3} \alpha^\sigma (Z \alpha^\sigma)) + Z \hat{O}_{2} \alpha^\sigma (Z \alpha^\sigma - 4Z \hat{O}_{\alpha^\sigma}) -
\]

\[
Z \alpha^\sigma \hat{O}_{\alpha^\sigma} 4Z \hat{O}_{\alpha^\sigma} 6Z \hat{O}_{2} \alpha^\sigma + Z \alpha^\sigma \hat{O}_{\alpha^\sigma} 4Z \hat{O}_{\alpha^\sigma} 6Z \hat{O}_{2} \alpha^\sigma 4Z \hat{O}_{3} \alpha^\sigma
\]

\[
\begin{vmatrix}
Z \alpha^\sigma & 4Z \hat{O}_{\alpha^\sigma} & 6Z \hat{O}_{2} \alpha^\sigma \\
1 & Z \alpha^\sigma & 3Z \hat{O}_{\alpha^\sigma} \\
0 & 1 & Z \alpha^\sigma \\
\end{vmatrix}
\]

\[
+Z \alpha
\]

\[
\begin{vmatrix}
Z \alpha^\sigma & 4Z \hat{O}_{\alpha^\sigma} & 6Z \hat{O}_{2} \alpha^\sigma & 4Z \hat{O}_{3} \alpha^\sigma \\
1 & Z \alpha^\sigma & 3Z \hat{O}_{\alpha^\sigma} & 3Z \hat{O}_{2} \alpha^\sigma & 4Z \hat{O}_{3} \alpha^\sigma \\
0 & 1 & Z \alpha^\sigma & 2Z \hat{O}_{\alpha^\sigma} \\
0 & 0 & 1 & Z \alpha^\sigma \\
\end{vmatrix}
\]

The first three terms of the determinant can be evaluated easily as

\[
Z^3 \alpha^\sigma \hat{O}_{\alpha^\sigma} \alpha^\sigma + 6Z \hat{O}_{2} \alpha^\sigma - 4Z \hat{O}_{\alpha^\sigma} (Z \alpha^\sigma - 3Z \hat{O}_{\alpha^\sigma} (Z \alpha^\sigma))
\]

\[
(1.4)
\]

whereas the last term of the determinant can be calculated using the same cofactor approach as above, which yields

\[
-4Z \hat{O}_{3} \alpha^\sigma + 3Z \hat{O}_{2} \alpha^\sigma (Z \alpha^\sigma) - 2Z \hat{O}_{\alpha^\sigma} (Z^2 \alpha^\sigma \alpha^\sigma) + Z \alpha^\sigma (Z^3 \alpha^\sigma \alpha^\sigma \alpha^\sigma + 6Z \hat{O}_{2} \alpha^\sigma
\]

\[
-4Z \hat{O}_{\alpha^\sigma} (Z \alpha^\sigma) - 3Z \hat{O}_{\alpha^\sigma} (Z \alpha^\sigma))
\]

\[
(1.5)
\]

Therefore, the function $f_5^L (-Z)$ as defined in Eq. (7.18) which is equivalent to $|M_1|$ without the factor $-Z \beta^{-\sigma}$ is given by

\[
f_5^L (-Z) = Z^5 \alpha^\sigma \hat{O}_{\alpha^\sigma} \alpha^\sigma \alpha^\sigma \alpha^\sigma + \hat{O}_4 \alpha^\sigma Z - \hat{O}_3 \alpha^\sigma (5Z^2 \alpha^\sigma)
\]

\[
+\hat{O}_2 \alpha^\sigma (10Z^3 \alpha^\sigma \alpha^\sigma - 10Z^2 \hat{O}_{\alpha^\sigma} \alpha^\sigma) + \hat{O}_2 \alpha^\sigma (10Z^4 \alpha^\sigma \alpha^\sigma \alpha^\sigma + 15Z^3 \alpha^\sigma \hat{O}_{\alpha^\sigma}) \hat{O}_{\alpha^\sigma}
\]

\[
(1.6)
\]

A similar procedure follows for the determination of the determinant $|M_2|$. The corresponding function $f_4^L$ is given by

\[
f_4^L (-Z) = Z \hat{O}_3 \alpha^\sigma - \hat{O}_2 \alpha^\sigma (4Z^2 \alpha^\sigma) + \hat{O}_\alpha^\sigma (6Z^3 \alpha^\sigma \alpha^\sigma - 3Z^2 \hat{O}_{\alpha^\sigma}) - Z^4 \alpha^\sigma \alpha^\sigma \alpha^\sigma \alpha^\sigma
\]

\[
(1.7)
\]
The equations of \( f^\sigma_2, f^\sigma_3, f^\sigma_4 \) can be determined similarly as shown below.

\[
\begin{align*}
f^\sigma_3(-Z) &= Z\hat{\alpha}^2 - 3Z^2\hat{\alpha}^\sigma + Z^3\hat{\alpha}^\sigma, \\
f^\sigma_2(-Z) &= Z\hat{\alpha}^\sigma - Z^2\hat{\alpha}^\sigma, \\
f^\sigma_1(-Z) &= Z\hat{\alpha}^\sigma, \\
\end{align*}
\]

(I.8) (I.9) (I.10)

Now, we can substitute these equations to Eq. (7.23) and determine the first five order commutations of \( S \) with \( H_V \). The first three order results can be used to verify the corresponding results determined from an easier way shown in chapter 7 whereas the fourth order is a new result shown in the following.

\[
\begin{align*}
\alpha^\sigma(2\beta^{-\sigma}(8Z^2\beta^{-\sigma}\gamma^\sigma\gamma^\sigma + 44\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma + 8\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma - Z^3\gamma^\sigma + 3Z^2\beta^{-\sigma}\alpha^\sigma) \\
-6Z\beta^{-\sigma}\alpha^\sigma + 8\beta^{-\sigma}\alpha^\sigma) + Z\omega^\sigma \omega^\sigma (4\beta^{-\sigma}\beta^{-\sigma} + 16\beta^{-\sigma}\beta^{-\sigma} + 2Z^2\beta^{-\sigma}\gamma^\sigma) \\
-4Z\beta^{-\sigma}\beta^{-\sigma} + 4Z^2\omega^\sigma \omega^\sigma - 12Z\beta^{-\sigma}\gamma^\sigma + 12Z\beta^{-\sigma}\gamma^\sigma + 6Z^2\omega^\sigma \omega^\sigma \\
+6Z^2\omega^\sigma \omega^\sigma + 2Z^2\alpha^\sigma + 3Z^2\alpha^{-\sigma} + 3Z^2\alpha^{-\sigma})] \\
\end{align*}
\]

(I.11)

The fifth order is similarly,

\[
\begin{align*}
2\beta^\sigma(4\gamma^\sigma\alpha^\sigma\alpha^\sigma\alpha^\sigma\alpha^\sigma(4Z^2 + 12Z^2 + 16Z + 16) \\
+2Z^3\beta^{-\sigma}\gamma^\sigma\alpha^\sigma\alpha^\sigma + \beta^{-\sigma}\alpha^\sigma\alpha^\sigma\alpha^\sigma\alpha^\sigma (96Z^2 - 96Z + 96) - 8Z^2\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma \gamma^\sigma \gamma^\sigma \\
-120Z\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\alpha^\sigma - 48\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma \gamma^\sigma + 40Z\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma \gamma^\sigma \\
+16\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma) \\
+\omega^\sigma \omega^\sigma ((32Z^2 + 60Z)\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma + (32Z^2 - 96Z^2 + 96Z)\beta^{-\sigma}\beta^{-\sigma}\alpha^\sigma \alpha^\sigma \\
+42Z^4\beta^{-\sigma}\gamma^\sigma \omega^\sigma \omega^\sigma - 12Z^3\beta^{-\sigma}\beta^{-\sigma}\alpha^\sigma \alpha^\sigma + (4Z^3 - 24Z^2)\beta^{-\sigma}\beta^{-\sigma}\omega^\sigma \omega^\sigma \\
+20Z^4\beta^{-\sigma}\gamma^\sigma \alpha^\sigma \alpha^\sigma + 32Z^5\alpha^\sigma \alpha^\sigma \omega^\sigma \omega^\sigma + 52Z^4\beta^{-\sigma}\gamma^\sigma \alpha^{-\sigma} \alpha^\sigma \\
-40Z^3\beta^{-\sigma}\beta^{-\sigma}\gamma^\sigma \alpha^{-\sigma} + 44Z^5\alpha^{-\sigma}\alpha^\sigma \omega^\sigma \omega^\sigma + 16Z^5\omega^\sigma \omega^\sigma \omega^\sigma \omega^\sigma \\
+16Z\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma}\beta^{-\sigma} + 16Z^4\beta^{-\sigma}\beta^{-\sigma}\alpha^\sigma \alpha^\sigma - 16Z^3\beta^{-\sigma}\gamma^\sigma \omega^{-\sigma} \omega^\sigma \\
+12Z^4\alpha^{-\sigma} \alpha^\sigma \omega^{-\sigma} \omega^\sigma + 2Z^5\alpha^\sigma \alpha^\sigma \alpha^\sigma + 8Z^5\alpha^{-\sigma} \alpha^\sigma \alpha^\sigma \\
+6Z^5\alpha^{-\sigma} \alpha^{-\sigma} \alpha^\sigma + 24Z^3\beta^{-\sigma}\gamma^\sigma \gamma^\sigma), \\
\end{align*}
\]

(I.12)
References


References


